

# DOCUMENT RESUME

ED 106 003

PS 007 998

**AUTHOR** Raizen, Senta; And Others  
**TITLE** Appendixes to Design for a National Evaluation of Social Competence in Head Start Children. Report No. R-1557-HEW.  
**INSTITUTION** Rand Corp., Santa Monica, Calif.  
**SPONS AGENCY** Office of Child Development (DHEW), Washington, D.C.  
**REPORT NO** R-1557-HEW  
**PUB DATE** Nov 74  
**NOTE** 216p.; For the complete Rand report, see PS 007 880; Pages 15 through 83 have been filmed from best available copy

**EDRS PRICE** MF-\$0.76 HC-\$10.78 PLUS POSTAGE  
**DESCRIPTORS** Classroom Environment; Cognitive Development; Curriculum; \*Ethnic Groups; \*Guidelines; Interpersonal Competence; Language Development; Minority Groups; \*Negro Attitudes; Preschool Education; \*Preschool Evaluation; \*Research Design; Social Development; Spanish Americans; Spanish Speaking; Standardized Tests; Teacher Background; Test Validity  
**IDENTIFIERS** \*Project Head Start

## ABSTRACT

This document contains appendixes to the Rand report which include: (1) recommendations for program evaluation and research made by a panel of 12 black professionals, (2) reaction papers by three black professionals and three Spanish-surnamed professionals, and (3) an expanded discussion of the technical measures proposed in the main report. Major recommendations and reactions focus on the definition of the term "social competence" and the independent variables involved, research design problems, projected outcome validity, and implications for black and Spanish-speaking children. A survey of literature concerning intelligence measurement and linguistic competence in Mexican-American populations is included. Technical information is given on tests used to measure perceptual-motor/cognitive/language skills. Information is included which supplements the discussions of social competence in Chapter 5 of the main report, particularly focusing on instrument recommendations in the areas of direct observation, evaluative responses from others, measures collected from subjects, and measures restricted to subsample studies. Also included are a literature survey on the independent variables, categorization of counties according to metropolitan/sparseness dimension, and an approximation of costs for basic battery testing per site. (ED)

ED106003

WN-8882-HEW

November 1974

U.S. DEPARTMENT OF HEALTH,  
EDUCATION & WELFARE  
NATIONAL INSTITUTE OF  
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-  
DUCED EXACTLY AS RECEIVED FROM  
THE PERSON OR ORGANIZATION ORIGIN-  
ATING IT. POINTS OF VIEW OR OPINIONS  
STATED DO NOT NECESSARILY REPRESENT  
OFFICIAL NATIONAL INSTITUTE OF  
EDUCATION POSITION OR POLICY.

APPENDIXES TO DESIGN FOR A NATIONAL EVALUATION OF SOCIAL COMPETENCE  
IN HEAD START CHILDREN (R-1557-HEW)

Senta Raizen, Sue Berryman Bobrow, with Tora Kay Bikson, John A. Butler,  
Karen Heald, and Joan Ratteray

A WORKING NOTE

prepared for the

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

*This Note is intended only to transmit preliminary research results to a Rand sponsor and may not be distributed without the approval of that sponsor. Views or conclusions expressed herein may be tentative and do not necessarily represent the opinion of the sponsor.*

PS 007998

**Rand**  
SANTA MONICA, CA. 90406

PREFACE

This document contains appendices to the Rand report *Design for a National Evaluation of Social Competence in Head Start Children* (R-1557-HEW, By Senta Raizen and Sue Berryman Bobrow, with Tora Kay Bikson, John A. Butler, Karen Heald, and Joan Ratteray) considered primarily of interest to the Office of Child Development or to a contractor who would manage the evaluation. They are not being made available by The Rand Corporation to a general readership. The document begins with Appendix C because Appendix A (list of panel members and consultants) and Appendix B (abstract of the Black and Spanish-speaking panels' responses) appear in the main report.

## **Appendix C**

### **CONTRIBUTIONS MADE BY PANELS OF BLACK PROFESSIONALS**

This appendix is composed of

- o A packet of final recommendations made by twelve Black professionals convened in a panel at The Rand Corporation in March 1974.
- o Three reaction papers (to Rand's Interim Report) prepared by the Black professionals who convened at Rand in January 1974.

### **RECOMMENDATIONS**

#### **INTRODUCTION**

Assessment by its very nature is a political act. It is political in terms of its potential impact upon the context where assessment occurs and in terms of the use to which the results of an assessment are placed. There is an impact upon assessor and assessed alike in terms of how both will feel about information presented and how that information will affect decisionmaking at all levels. The very act of assessing defines a power relationship. Someone is seen as "able" to assess and someone else as "object" of assessment. Therefore, independent of the activities within the process of assessment, the consequences of doing an assessment program can, if improperly handled, become a part of the fabric of a system which reinforces in minority communities a sense of powerlessness regarding factors which control their lives.

We are justifiably suspicious and anxious about research and testing involving Black people because of the history of such research activity at the national level and the damaging consequences of distorted, often incompetent research and measurement. The distortion

and incompetence are often present because of several things. First distortion occurs because the "state of the art" in the measurement of complex human behavior is embryonic to say the least. And yet, once results are obtained, they tend to be treated as facts--especially if the data came from "reputable" sources. Distortion also occurs because of careless, shoddy, and irresponsible implementation of the best standards which we have even given the embryonic state of the art. Research and measurement often occurs and is seen as competent because of the reputation of organizations or institutions which conduct and support such research and measurement. What is absolutely required, however, is that competent research work involve knowledgeable researchers who understand the process of education intimately. It must be clearly understood that competence in one area of social science or science does not necessarily transfer to others. Engineers, economists, and research sociologists alone may be "interdisciplinary" but cannot be considered competent to assess schools until the qualified educators in the discipline and practice of education are included.

The failure to take such things as the above completely into account yields studies such as the Coleman Report. We have seen how this study which covered only a few faulty and in most cases relatively irrelevant variables and which used a highly questionable experimental design has now become the foundation among many policymakers for decisionmaking regarding the "effectiveness of schools," and the financial support for schools. Black children and their families have suffered from the tendency of policymakers to accept these findings prematurely. As a result, any similar evaluations involving Black people must operate only under stringent mechanisms for quality control. Too much is at stake to permit evaluators to take short cuts or to make compromises when assessing the Head Start program.

## GUIDELINES FOR ESTABLISHING ACCEPTABLE RESEARCH DESIGNS

### Design Specifications

1. Control of significant environmental variables, e.g., population distribution, health-delivery systems, food distribution as they relate to nutrition and health.
2. Concise and parsimonious research design.
3. The use of pilot tests in nationwide evaluation research should be avoided. Such pilot testing should be limited to few sites and should be separate from the overall evaluation procedure.
4. Exploratory variable measures should not be included in evaluative design.
5. There should be a clear explication of the contractor's scheme and theoretical base (i.e., the use of principled research personnel who have demonstrated skill in educational/psychological research, classroom instruction and appreciation for relevant community variables. They should also have specific experiences in early childhood education and research.
6. Use of multi-disciplinary theoretical approaches (e.g., consideration of economic, political, sociological, and anthropological theory and data).
7. Test procedures and results should facilitate local purposes.

### Black Community/Professional Involvement

1. Black professional groups (e.g., Association of Black Psychologists, etc.) should serve as monitors of research teams, i.e., contracted

research designers, implementors and interpreters.

2. Research teams at all levels should be multi-ethnic, functional and acceptable to the above monitoring groups.
3. Involvement of local (at test site) review panels composed of professionals, parents and Head Start staffs. H.E.W. guidelines or informed consent for research on human subjects should be utilized.
4. Feedback of results to service-recipients.

#### Test Criteria

1. Professional standards of validity and reliability should be adhered to (e.g., use of A.P.A. technical recommendations for psychological tests and diagnostic techniques). Standards of minority professional groups must also be adhered to.
2. Use of in-depth and longitudinal assessment methods as opposed to superficial assessment devices implemented for expedience. It is also necessary to utilize qualitative culturally relevant measures.
3. Measures should be specific to relevant operationally defined variables.
4. Items selected from previously validated test package must be singularly validated prior to reuse.

RECOMMENDATION I

The proposed sample size -- 150 classes -- comprising two Head Start cohorts is unrealistic for a longitudinal study. Even though Rand's latest shift (March 30, 1974) indicates that it has been decided to follow the two cohorts for only one (1) year, the number of important independent variables by which the sample cohorts would have to be stratified would realistically require approximately 960 classes and appropriate comparison children initially. A mechanism must be devised to identify and follow these children; many who will move within the course of the study.

The term comparison group should be used rather than control group since the selection of a control group in an experimental sense is likely to be difficult.

INDEPENDENT VARIABLES

Parent/Child Ethnicity: (Black American - Other. e.g., Caribbean)  
(8) Chinese/Oriental  
Indian  
Italian  
Jewish  
Spanish Speaking Caribbean - Other  
Mexican American  
White (Other)

Region:  
(6) Northeast  
Southeast  
South Central  
North Central  
Southwest  
Northwest

Residence:  
(5) Rural - Farm, Migrant, Town  
Urban - City, Suburb



Program Type:

(21)

Public School System

Community Agency

$8 \times 6 \times 5 \times 2 = 480$  cells x (minimum) two classes = 960 classes

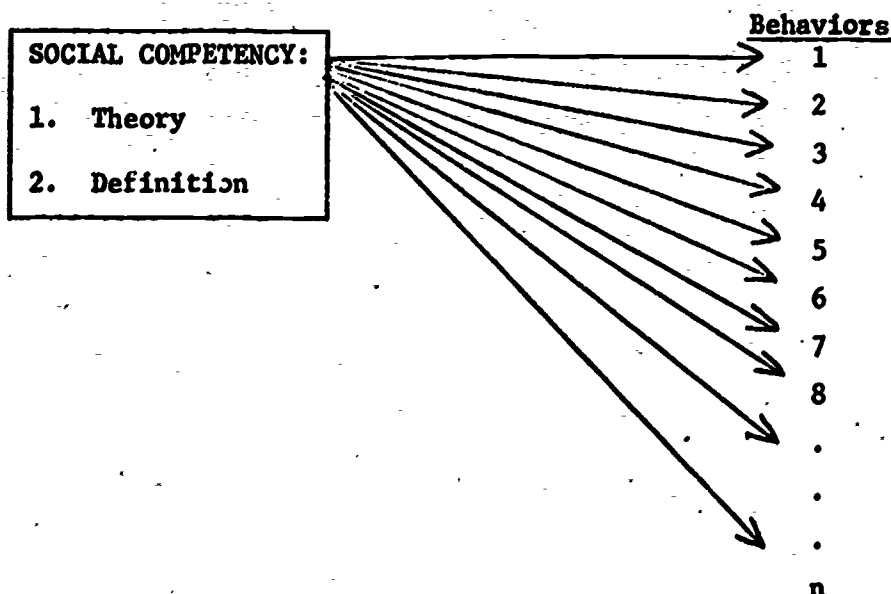
## RECOMMENDATION II

Social competence must be defined in such a way that it clearly delineates and delimits the behavioral domain to which the construct applies. The present global, all inclusive orientation or approach, presents certain methodological problems in selecting variables and measures of social competence.

Because of the absence of any theory and definition of social competence as the basis for selecting the relevant behaviors, it is not possible to construct an analytic framework which would guide research methodology and establish measurement priorities. Normally, the procedure followed by social scientists to construct behavioral measurement tools is to proceed from:

- (a) theory, to
- (b) operational definitions, to
- (c) specification of the behaviors that the definitions encompass, to
- (d) establishing priorities among the set of behaviors that would be measured, to
- (e) test construction and field testing for refinement of the instruments.

Each step in the procedure above involves apriori (value-laden) judgments. The paradigm is illustrated below:

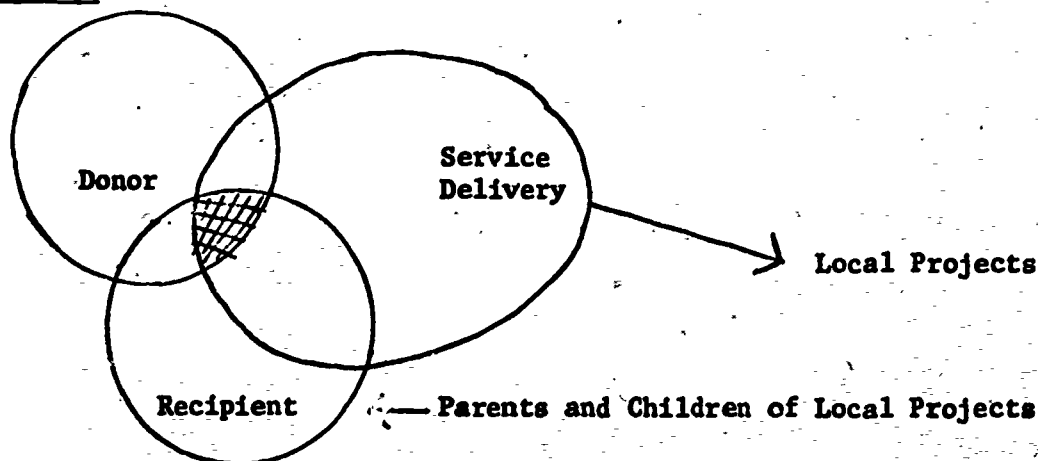


Given the various ethnic populations in Head Start, if this paradigm were used, one would find that each of the different groups has and would select specific different behaviors or combinations of behaviors for the delimited set as having high value to the group. For example, from the behaviors above that have been specified as being in accord with the theory and definition of social competency, Black communities might select B1, B2, and B3 as highly valued; Puerto Rican communities might select B1, P3, B4, B5; Mexican American communities might select B3, B5, and B6. However, no community can select the behaviors -- neither academic nor social -- that it values if all behaviors are designated by the Rand Corporation as being important.

Again, because of the lack of a definition of social competence that would permit identifying the behaviors that should be measured child outcome measures are confounded with general program input variables. In January 1974, Mr. Murphy stated that the programs as such were not to be evaluated. The panelists assumed that this was OCD's position. Nevertheless, the first and present panelists insisted that outcomes for children could not be realistically separated from the direct transmission of information, attitudes, and skills by adults to children in Head Start classrooms. The panelists were not referring to overall instructional program goals (as expressed by program directors or Head Start boards), but to the actual constituents of instruction as they occur in classrooms. The reason for excluding program goals in the evaluation design is that there are always discrepancies among national Head Start goals, regional goals, local program goals, teacher goals, and the interpretation of the goals by individuals. Examples of the confounding of program evaluation with child outcomes occur on pages 8-10 of the Interim Report summary, where health program service delivery variables -- B. "Absence of Illness/Impairment," C.3. Better health service use; page 3 of TEST DEVELOPMENT TASKS (March 30, 1974), where it is considered necessary to develop "... assessment techniques for health service use. In order to check the health care goals of Head Start...." and on page 2 of BASIC EVALUATION '75 (March 30, 1974), where "Planning/Supervision," and "Center Sponsorship" are listed as independent variables.

Any service delivery program must take into account the fact that the "program" consists of three sub-systems: Donor, Service Delivery, and Recipient.

Illustration:



Each system has its own value orientation about what the goals of a Head Start program should be.

An efficient evaluation design, therefore, must take into account the value orientations of each of these subsystems in identifying outcome behaviors which serve as indicators of social competence.

Because many categories of behavior lack clear specification and necessary detail, some behaviors have been included which have little or no positive relationship to the school behaviors specified, approved, and rewarded in the primary grades, viz. physical vigor. The primary grades value passivity, cleanliness and neatness, and verbal assertiveness. Parental involvement with the teacher can be helpful for success in school.

In other words, there is a lack of congruence between values and behaviors which are often the objectives of H.S. programs and the values and behaviors (objectives) of the primary grades. Consequently, the incongruence of H.S. and primary grade objectives often militate against high transferability of behaviors from H.S. to primary educational programs.

**RECOMMENDATION II.1: PROTECTION OF HUMAN SUBJECTS**

a) See H.E.W. guidelines

Re: Informed consent, confidentiality, etc.

In order to administer tests, not only should directors be informed of the items being used and the purpose of the tests, but parents, children, and the community must have a say about the acceptability of tests proposed. Consent must be given by parents and community before such an endeavor be undertaken.

Reasons for undertaking such a task should be explained by the individuals involved in testing.

b) No results from this study should go into local records.

RECOMMENDATION IV: REGARDING STANDARDIZED TESTING

To use a base battery of tests on a national scale is not a valid procedure because a base battery of tests would not take into consideration the different locales and ethnic groups. We would be perpetuating biases because what is a natural behavior for one group of people is unnatural for another group due to different cultural experiences.

OCD and the Rand Corporation must be aware that Head Start was not founded in a manner that would permit accurate, undistorted evaluations using standardized tests. Rather, it was founded on the principle of local autonomy in the socialization of young children supplemented with academic skill preparation.

Below are three fallacies which would be ignored if standardized tests were used:

1. There are no measurements that are not biased.
2. Most test items are geared toward middle class urban America.
3. There are no instruments that would measure any diversified group of children fairly.

In Rand's Report no consideration was given to the Vineland Social Competence Scale developed by Edgar Doll several decades ago. This scale was designed to measure middle class "white" American child behavior, but it can be modified on the basis of other ethnic group behavioral values. It is realized that Rand has taken into consideration five areas of development, namely, Health and Nutrition, Perception and Cognition, Language, and Social and Personal Development, and utilized subtests from a number of sources. The Vineland covers all of these areas and gives consideration to local variations and could be adapted for minority groups.

If a base battery is used across all cohorts and ages to test academic readiness or achievement, it should only contain skill-specific tests for, say, reading, mathematics, understanding of physical principles, etc.

In summary, it is clear to the panelists that although the Corporation has been "working" on the evaluation design for some eight months, much of

the basic and essential preparatory research prerequisite to the design has not been done. Evidence for this position is explicit in the several changes in the content of the evaluation in response to Black and Mexican-American professional information and pressures. This process should have been initiated early in the work rather than three months before the final report must be submitted. Had that been done, OCD and the Corporation would now have a scientifically acceptable definition of social competency, comprehensive enough to apply to all of the ethnic populations in Head Start, and the evaluation design would probably be much different from what it presently is.

**RECOMMENDATION V: PREPARATORY ACTIVITIES FOR EVALUATION DESIGN AND SAMPLING**

Although it would appear that the "best instances" of the Head Start treatment by OCD and national Head Start criteria, in order to build in some uniformity for a crucial independent variable, it is suggested (recommended?) that programs be selected randomly; that the kind and quality of program services be recorded and categorized across programs in the same or different locations with similar ethnic populations; and that the program services categories be used as a multiple set of independent variables that will serve to separate different treatment groups for data analyses. This procedure would eliminate the probability that OCD, national Head Start, or the Corporation will superficially select "best instances of treatment" and also avoids the necessity to choose among slightly different "treatments" in the same regions and populations, although all of the programs conform to basic Head Start implementation criteria.



BLACK PANEL'S STATEMENT OF ITS PROFESSIONAL RESPONSIBILITY

The work of the Black Panel convened by Rand on March 29-30, 1974 is not to be considered as an explicit or implicit endorsement of Rand's work done in the past -- or currently -- on the design of the Head Start Evaluation. It is to be understood that when the evaluation design is completed, that the Black Panel will be re-convened. Failure to do so will result in automatic disclaimer by the Black Panel Group. The reconvening of the Black Group will not imply endorsement. That meeting will provide an opportunity for review of the design to see if the recommendations have been followed to the extent that the design is considerably altered so that it reflects the input we feel is necessary for an effective job.

March 30, 1974

**CRITIQUE AND COMMENTARY ON THE INTERIM REPORT**

**BY THE RAND CORPORATION:**

**"DESIGNING AN EVALUATION OF SOCIAL COMPETENCY  
IN HEAD START CHILDREN"**

**Patricia E. Allen  
Black Child Development Institute  
January, 1974**

OVERVIEW OF THE POSITION OF THE BLACK CHILD DEVELOPMENT INSTITUTE  
IN REFERENCE TO THE RAND EVALUATION DESIGN

"We believe that the developmental years of life are crucial to the maturation of children. Life stresses placed on families in this society have raised the need for comprehensive child care to an extremely high level." It is a premise of the Institute that a child must first function within the context of his family and his community.

It is the position of BCDI that the most appropriate unit of government to relate to these concerns is the Office of Child Development of the Department of Health, Education and Welfare whose responsibility it is to facilitate the development of a program which will meet individual needs of Black children as well as the majority group needs.

As the only national Black child advocacy organization, the Black Child Development Institute has a responsibility to monitor, analyze and share with the Black community the activities of the Office of Child Development, as they relate to Black children and their families.

We insist that:

1. Services must impact concretely and positively upon the lives of children;
2. Comprehensive child development services must take place in the context of overall community development; and
3. Services must reach children through the active and decisive participation of their parents and community leaders.

It is from this viewpoint that we draw the following analysis between the Interim Report of the Rand Corporation and previous studies conducted by the Office of Child Development.

BCDI in its review of Office of Child Development program initiatives such as the Child Development Associate, the Model Licensing Codes, and Head Start Improvement and Innovation continues to identify a faulty planning process. Moreover, each of these programs seem to have as its rationale cost-cutting concomitant with inadequate emphasis on quality child care. The result of this faulty planning process leads to programs that are either ineffective or actually harmful to children.

BCDI takes this opportunity to again recommend to the Office of Child Development that any program development or evaluation design should in the beginning stages:

1. Respect the authority of parents and utilize their expertise at all kinds of decision making.
2. Capitalize on the experience and knowledge of Head Start operators and personnel at the local level.
3. Be consistent with common rules of data collection and analysis to insure reliable data.
4. Involve Black scholars in research efforts that will impact the lives of Black children.

In conclusion we wish to point out that it is meaningless to call in a group of Black consultants when there is nothing left to consult about. It is our position that if this group is to have an input which will restructure the design of the evaluation of social competency it will be necessary to have a commitment from the Office of Child Development to include their recommendations in the final design.

THE RAND INTERIM REPORT  
January 25, 1974

In response to the request to analytically review the Rand Interim Report On Designing An Evaluation Of Social Competency In Head Start Children, with particular reference to the proposed measures and their potential effect on Black children who, according to the 1970 census, make up 50% of the Head Start population, the following report is submitted:

After repeated analysis of the interim report, we have decided against a line by line critique resulting from our examination. It is our position that the technical design needs to be revised. This procedure will require extensive input over a period of time. Proceeding on the premise that the Rand Corporation has contracted on July, 1973, to design an instrument to be used to evaluate the social competency of Head Start children our first concern is that there must be an assurance that the characteristics to be evaluated can in fact be measured and that the proper instrument is available or will be constructed to measure social competency? It is our considered judgment that such an instrument can be developed provided it is approached from a comprehensive perspective.

In order to achieve the above, it appears logical to proceed in the following manner.

There is a need for Office of Child Development and Rand Corporation to examine the current evaluation study of the Performance Standards. This would provide for a process whereby there would be established a clear and consistent perception of the meaning of the Performance Standards.

In order that the objectives of this report be realized it will be necessary that OCD and the Rand Corporation re-examine and re-define the contract specifications. We are recommending that prior to the finalization of the contract there should be collaboration with experts in the field of Early Childhood Education and

parent organizations from varied ethnic backgrounds. In order to develop and implement a valid measurement of social competency in Black children it is mandatory to have continuous input from a Black researcher or expert who has extensive experience in Early Childhood Education.

Site visits to centers in the various regions are necessary to make possible the use of an experimental approach in the development of the evaluation design by Rand personnel.

The interim report is grossly incomplete as a technical instrument which is to be applied for field testing in September, 1974. It is our considered opinion that it will need extensive restructuring before it should be field tested. Consideration must also be given to the fact that the field test must be administered in September of 1974 to comply with the renegotiated time table. Another aspect to be considered are the drastic program changes which are presently underway in order to comply with the Performance Standard. The above facts must be taken into account in establishing a realistic timetable.

The interim report did not deal with a specified procedure for Pre and Post testing. If the design is intended to accurately measure the social competency of Head Start children a procedure for controls must be established to insure the attainment of the objectives of the evaluation.

As a follow up to the January 21st and 22nd meetings the consultants are to:

1. Submit to the Rand Corporation by January 29, 1974, the names of approximately twelve Blacks with expertise in the various disciplines of Child Development, for the purpose of critiquing the revised interim report as suggested by the consultants.
2. Confer with Ms. Senta A. Raizen to establishing the agenda for the March 4 meeting.
3. Receive from the Rand Corporation, copies of relevant excerpts of the interim report by February 1.

4. Receive critiques of the proposal by February 15th. Study critiques and make recommendations to the Rand Corporation.
5. Submit names of six persons to be invited to a two day meeting the week of March 4. The Black consultants will meet separately the first day to review the agenda, discuss the information contained in the critiques and develop a plan of action. On the second day the original consultants, the invited experts and the representatives of the Rand Corporation will meet jointly.

Patricia E. Allen  
Curriculum Specialist  
Black Child Development Institute

**Critique and Comments on the Technical Proposal  
and the  
Interim Report**

**Designing an Evaluation of Social Competence in  
Head Start Children**

**The Rand Corporation, Washington, D.C.**

**Harold Freeman, Jr.**

**January 1974**



## TECHNICAL PROPOSAL

### Designing An Evaluation Of Social Competency In Head Start Children

The evaluation design work plan touches most, if not all, bases. However, two areas that are not discussed (which does not necessarily mean that they won't be taken into account) are: (a) the basic criteria or standards that will orient the search for written materials and the selection of "standardized" instruments, or those that will be developed, and (b) generally, how differences between rural and urban Head Start children will be handled; whether they will be taken into account in selecting assessment tools, control or comparison groups, data analyses, and so forth.

In the final analyses, the most critical variable for both areas will be the "experts" selected for the panels and the research personnel who will implement the evaluation, with respect to their orientation towards poor, Black and white children and adults and community Action in rural and urban environments; if the vast majority of the "experts" are white traditional academicians, they will, as usual, apply their values and standards - their set and experiences - in proposing criteria and metrics for what is, in effect, other groups with different values and standards for behavior. Having a few fully assimilated "Black" folk on the panels will not make a substantial difference in the outcome of the deliberations and recommendations.

-43-  
INTERIM REPORT

Outcomes And Their Assessment

Page 2:

"...other factors that enable a child...": social competence is not adequately defined; though it is possible to transform the general concept into specific behaviors for children from different ethnic Head Start populations, that has clearly not been done. But if it isn't, the proposed work has no merit and does not deserve funding because the metrics that are being considered for inclusion in the evaluation are the same ones that have been employed in the past in small-and large-scale studies and they will yield the same findings.

Page 11:

It is becoming more apparent to students of child development that the psychological measures and social indices constructed for all children are, in fact, based on a narrow range of experiencing and an object-oriented set of values representative of the majority culture's perception of the ideal behaviors prerequisite for optimal functioning in the American socio-economic system. Those "minority" groups who are not a part of "mainstream" America, whose reality orientations are different, whose social/personal skills are different in content and range, and, finally, who are not taught the same role outcome expectations have never been properly evaluated. The content of present measures do not take into account any minority culture psychological styles or value combinations. Therefore, the minority child's performance on tests based on majority culture values and ideal behaviors, is a measure of the degree to which the child's parents and teachers have taught him/her to conform or to comply with the majority culture's definition

of the "good, intelligent" child. The standardized, normalized tests, then are in fact achievement or criterion-referenced tests for majority culture children and they are used as predictors of later competence for those children. Thus, the case against criterion-referenced tests for other ethnic groups is invalid and group-specific criteria for some of the five areas to be assessed have to be developed if one wishes to determine how well the children are functioning. This is especially true for cognition, language, socio-emotional, and motor/perceptual competencies.

Page 12:

A search of the literature relevant to and descriptive of Black children will yield as much theory as one ever needs. That has simply not been done to lay the foundation for this kind of work (see, for example, St. Clair Drake, W.E.B. DuBois, E. Liebow, A. Billingsley, K. Clark, E. Labov, L. Turner and W. Stewart).

Page 25:

It is inane to even suggest a survey of "eating habits" since only in extreme cases are there clear relationships between frequency and amount of food-ingested and general health. One cannot put a positive value on "three or four squares a day" vs. "one or two squares a day" since there is considerable variation between and within cultures here and elsewhere with no apparent negative effects on health.

Page 36ff.:

Language Development

The semantic component of verbal or non-verbal behavior refers to meaning - not to vocabulary! This error also appears in the summary of Featherstone's papers.

Page 39:

"...progress in standard English" is only important if there is a demonstrated communication problem, i.e., where teachers and children do not understand each other well, and the teachers model standard English; or where the teacher models standard English, expects the children to imitate her, and punishes those who don't. In the majority of the Head Start classes I've observed in, the teachers and children understand each other very well; and whether the teachers are White or Black, they only correct certain technical kinds of language behavior (e.g., grammar, semantics, vocabulary), and leave style (e.g., use of markers, intonation, omission of plural endings) alone because the children's and teachers' styles are often either in accord or are not barriers to effective communication. There is no evidence (p.37) that an English dialect for cognition is any less effective than standard English for cognition.

Page 41:

What evidence is there for the association between phoneme discrimination and reading skill level?

#3: If the PSI is the one Betty Caldwell developed; it is clearly majority culture bound and, for that reason, N.Y.C. Head Start refused to permit its administration there.

Page 42:

Was the validity and reliability of the ETS CIRCUS battery been established for Black Children?

There is in the language assessment section, no recognition of the unique dialects that Head Start Children throughout the country speak. The emphasis on standard English as the standard by which the children's linguistic competency will be assessed will actively mask their competencies.

Page 45:

Cognitive Development

The word "cognition" covers all of the various modes of knowing, perceiving, remembering, imagining, conceiving, judging, reasoning, and is normally contrasted with the affective (feeling) and conative (willing) modes or aspects of conscious life. Its use in American psychology, generally, and in this report is restricted to specific positively valued majority culture skills, e.g., mathematics, French, (p.48), reading, that are not uniformly taught to all children in the same way or for the same end.

Page 103:

The assumption that "social class differences in receptive vocabulary will put lower-class children at a communicative disadvantage" is based on the use and acceptance of standard English as the standard for linguistic production and reception. The fact that communication among Head Start children and teachers who speak dialect variants of standard English suffer no communicative disadvantages is not taken into account. But, if it isn't, both linguistic competency, specifically, and social competency, in general, will not truly be assessed for these populations.

Page 104:

Basil Bernstein's findings and interpretation of language behavior into the dichotomy "elaborated" and "restricted" coding for children in London, England has not been shown to be generally applicable to American White children. It is doubtful that his scheme has any applicability to American Black, Chicano, Indian and Chinese Head Start children. The extent to which Bernstein's data and interpretation for another culture is appropriate for dialects more or less distant syntactically and semantically from American Standard English is not known.

The place where language is used is critical for many children. Questioning, narration, etc., may occur in one place (e.g., at home) and much less so in another place (e.g., at school). How much or little different kinds of language behavior occurs, and how much of it occurs is very often a direct function of the expectations of the teachers and parents in the childrens' immediate environment. If school behavior will be the focus of the proposed research, it will be essential to know how much Head Start teachers encourage and reinforce language behavior in different populations of children. One cannot blandly assume that language use is equally reinforced among Black, Chicano, White, Chinese, etc. children, because adults from these groups place a different value on verbal behavior in children.

Page 128:

"What important outcome variables have been omitted?"

1. Children's perception of and attitudes towards (a) elementary school teachers, (b) school or school-like environment, and (c) the content of elementary school curricula in grades K through 3.
2. Measures of test and general anxiety (see, e.g., Sarason, et al., Castenada).

3. Children's perception of parent and teacher potency, power in the context of the (local) educational establishment.
4. Head Start, non-Head Start parental modelling effects on children as a function of parent participation in preschool educational programs.
5. Elementary school teacher performance expectations for Head Start, non-Head Start children.

#### SUMMARY

Although the evaluation is supposedly aimed at measuring the "social competence" of Head Start children, the phrase "social competence" is not sufficiently defined to permit its translation into specific, measurable social behaviors that particular groups of Head Start children learn. Instead of defining precisely what social competency (i.e., general environmental coping skills) means and delimiting the relevant types of behavior it could include "everything" that children can do - all of their physical functions (but not all psychological) were included in the definition, and in the absence of measures appropriate to the research focus, traditional psychological and language tools will be used. It is not possible, under these circumstances, to get data different from what has already been collected by others - though the interpretation (i.e., the conclusions) may be different in this case.

If social competency refers to coping skills, then clearly the combinations of such skills taught to different ethnic groups are different as a function of the kind and quality of environmental press - that is whether, for example, the personal environment in which children develop is permissive, supportive, destructive, punitive, hostile, etc. Nowhere in the report is there the



slightest hint that the metacognitive information and skills (see page 52) by which each ethnic group or subculture organizes its responses to the environment was seriously considered. There is, in fact, no evidence in the report that such skills have been enumerated for white children, much less for Black and other ethnic groups who constitute the majority of the Head Start population.

Up to this point, the evaluation design work is more fantasy than reality, if social competency is truly its purpose. Most of the traditional measures that have been tentatively selected should indeed be used. But, in light of their bias they cannot begin to describe adequately the competencies of Head Start children. At least half of the final measures should be devoted to specifying ethnic group - specific inputs to the children as a preliminary to an enumeration of the important metacognitive skills that should be assessed. The kinds of skills, and their incidence or frequency in the Head Start children should be correlated with their performances on majority culture "cognitive" and "language" measures.

The health and nutrition and motor/perceptual assessment areas need little revision.



-30-  
**SOME IMPORTANT STRATIFICATION  
(Independent) Variables**

**Parent/Child Ethnicity:**

Black American - other, e.g., Carribbean  
Chinese/Oriental  
Indian  
Italian  
Jewish  
Spanish-speaking Carribbean, other  
Mexican-American  
White (other)

(Only some groups may be selected for the final samples, for some measures.)

**Region:**

Northeast  
Southeast  
South Central  
North Central  
Southwest  
Northwest

**Residences:**

Rural - Farm, Migrant, Town  
Urban - City, suburb

**Program Type:**

Local Public School Systems  
Community Action Agencies

Teacher Ethnicity in combination with parent/child ethnicity (important because adults from different ethnic groups transmit some different coping skills differentially to in-out group children in their care.)

**Child Age**

**Grade in school (for post Head Start children)**

**Class within Grade**

**Socio-Economic Status**

**Presence/absence of Special Programs in elementary school (e.g., Follow Through**

CRITIQUE AND COMMENTARY ON  
DESIGNING AN EVALUATION OF SOCIAL COMPETENCY IN HEAD START CHILDREN

Arvern Moore  
Head Start Program  
ICS, Inc.  
Holly Springs, Mississippi  
January, 1974

I have reviewed the purpose and scope of the Rand Task. According to my understanding you are to develop an evaluation design to measure the social competence of Head Start children as defined by OCD in notice 30 364-1. Based on my review of the OCD and Rand definition of social competence it means to improve certain skills of disadvantaged children. You have broadly defined social competence to take into account the cognitive-intellectual, physical and mental development, and health and nutritional needs.

For the most part, a quantitative and objective measurement of these competencies, in my opinion, would be based on subjective data and findings. To design an evaluation instrument that encompasses the total Head Start operation and uses subjective, hypothetical, and assumptive data is an injustice to the accomplishments of Head Start.

On the other hand, if we are to truly do justice to Head Start it would seem only fair that in identifying those competencies that are necessary for a child to gain in order to function desirably in our society, all of the skills or benefits either learned or gained by the child while in Head Start should be included in the list of social competencies. Further, justice cannot be done to Head Start in an evaluation design that does not take into account the social competencies gained by the parents and the subsequent impact on other children in the home.

From the Head Start point of view it seems impossible to justify the assembling of panels to develop evaluation designs to measure any facet of Head Start without including on these panels parents of Head Start children and Head Start program operators. It is difficult to accept the fact that a fair evaluation design could be developed utilizing only researchers that, I would assume, have not had a large degree of exposure to Head Start on a local or community level.

### Socio-Emotional

I would further recommend that any evaluation design that could not be developed and based purely on empirical evidence should not be used to measure the impact Head Start has had on children or their families. Therefore, it is extremely difficult for me to accept an evaluation instrument whose basis is derived from theoretical assumptions.

In developing and identifying variables to be considered in an evaluation design, the variables considered and used in the implementation of the design should have some provision for controls. Without controlling the variables that could possibly influence the outcome, the assessment of the outcome, in my opinion, would not be reliable. Since apparently the identification of independent variables is difficult to isolate, I would suggest that the evaluation results would also be difficult to justify as accurate. The cost for developing and implementing a theoretical or hypothetical design is certain to be astronomical. With this known fact in mind, it would be my specific recommendation that since it is necessary for Head Start to be evaluated, a practical and realistic evaluation design be developed. By practical I mean a design that could be used by program operators, HEW, and contractors to evaluate Head Start. Also, in considering a design the scope should be broad. I feel this is necessary because by utilizing an evaluation design which is broad and general, then all of the impacts Head Start has had on the total community could be measured. This could be done only through an assessment instrument that was comprehensive enough to measure the outcome relative to the children, their families, and the total community. I feel that to limit the evaluation only to those children who are participating in Head Start is limiting the controls and will not give us the kind of identifiable outcome necessary for determining if children are reaching the overall

goals of Head Start.

Further, in my opinion, any evaluation design should take into consideration how far a child has moved, in what direction the child has moved, what standards are being used to identify how far the child has moved and what we consider ultimate as far as where we expect the child and his family to be at the end of his Head Start tenure. If the standard for this determination is to be norms that already exist it should take under consideration the different localities, cultural or ethnic groups, economical background, or environmental conditions. I contend that the results will not actually measure the effects of Head Start on the family and the child. Using existing norms as a standard would instead measure where the standard says the child should be and not where the child actually is at the end of his Head Start intervention and would not give due credit to where Head Start moved the child.

It is my opinion that any evaluation design which embodies as a standard the level that middle class children have achieved is not an acceptable instrument for Head Start. Further, since all Head Start children are not integrated into majorities, any evaluation design developed should consider the effect on children leaving Head Start and entering minorities. If control groups are to be used in this project, how would they be selected and how would you control the variables as they relate to the individual families from which the children come?

Head Start is based on the philosophy that each local community should design the program based on local communities needs. If the evaluation design does not take into account the needs of the local communities being served by the Head Start program, then I suggest to you that the design will be contrary to the purpose for which Head Start was created.

### Phase Validity of Outcome

I feel that one fundamental question should be resolved, that being the validity of the results gained from any evaluation instrument. If there is the slightest doubt in anyone's mind about the validity of the proposed instrument, the instrument itself should not be used because the end results, I feel, will certainly be detrimental to the accomplishments of Head Start. In addition, despite the fact that OCD and Rand have given their definition for social competence, it is extremely difficult for me to accept social competence as the overall goal of Head Start. This is true in the first place because, according to my understanding, the current performance standards are interim. This means that in the final writing of the performance standards, it is highly possible that social competence will not be the overall goal of Head Start. I am only urging that you do not become too presumptuous.

Another matter that concerns me is how will the test be administered and by whom will the test be administered. It has been our experience in Head Start that tests administered by faces unfamiliar to the children, either on a one to one basis or on a group basis, do not render the same results as tests administered by a familiar face on the same basis.

Approaching the instrument from the long range impact on children, if the design cannot be developed and implemented by covering a long range period, then I feel that the results will not adequately represent what the Head Start intervention has meant to the child and his family. Secondly, if necessary changes cannot be made in the school systems the Head Start children will be entering upon completing their Head Start tenure, then the long range impact of Head Start will certainly be adversely affected by the new school experience of the Head Start child.

In a program which is experimental or demonstrational itself, it is my opinion that experimental or unrealistic evaluation designs should not be used to determine the effectiveness of an experimental program. I would strongly urge that every outcome that can be measured quantitatively and without ambiguity should be included in any design that is used for the purpose of measuring the effectiveness of Head Start in improving the social competence of children and their families.

The evaluation design should be constructed and based on data with unquestionable validity and reliability. I suggest this not because of the fact that this kind of design would be simpler to develop, but because I feel that this kind of design would be more credible and would not do Head Start an injustice. Any other kind of design developed would certainly create national controversy and the validity of the data resulting from the design would be very suspect.

In considering outcomes that are important to determine "social competence" I fail to understand why priority groupings of outcomes are necessary. In looking at the total child it is my opinion that all outcomes are important in helping the child to gain those skills necessary for successful existence in society and should be weighed equally.

#### Health and Nutrition

Regarding health and nutrition, we feel in Head Start that some of the greatest measurable outcomes are accomplished through health and nutrition programs. With this being true, all outcomes related to health and nutrition should be included in the design. In addition, I feel that any design which fails to take into consideration the expectations of an uneducated school system with regard to what the Head Start intervention is supposed to accomplish, is defeating within itself. In other words, regardless of what the impact of Head Start is, an uncooperative and uneducated school system can very easily destroy it.



## Motor Perceptual

As it relates to health and the child's physical development, I don't feel that an evaluation of the Head Start effects would be valid and would not necessarily measure what effects Head Start has had on the motor development--either gross or fine motor skills. Readiness and developmental stages cannot be forced, therefore, age level expectations of primary teachers do not determine the stage or age of readiness of an individual child.

## Perceptual Processes

I feel that there is not a need to measure through a national test the perceptual development of Head Start children. This would not be valid since many of the tests used are in experimental stages.

## Language

With all of the cultural, ethnic, and geographic dialects and other variables that should be considered, I feel that a language test developed to evaluate Head Start would not give valid results.

To measure the effects of Head Start as it relates to linguistic competence would be impossible because this would be only measurable as to individual assessment and treatment from specialists. Head Start is not designed to do this.

Competence in language can only be determined by the objectives of local Head Start agencies and the needs and assessments of individual communities and outcome of the teaching process. As local program operators, we cannot assume what parents want for their children. In order to design programs on a local level that will meet the needs of the children and their families it is necessary that we get factual data from the parents, thereby providing them with an opportunity to design programs according to their children's specific needs. To develop programs contrary to this



would mean that our value judgments as to what we feel the parents need is taking precedence over the actual needs of the children as the parents see them.

### Cognitive Processes

To draw up a large scale evaluation of the impact of Head Start on cognitive development would be strictly judgmental and would do an injustice to Head Start since most tests designed to measure cognitive development are so institutionalized that they do not concern themselves with individual child development.

### School Readiness

National tests that have been used to determine school readiness have not been valid in measuring readiness skills of children in all areas. I feel that a national test would be difficult to develop and would be an injustice to Head Start in an overall evaluation. In addition, I recommend that consideration be given to evaluation designs that could determine if the school systems are ready for the children after their Head Start experience.

There are too many judgmental values related to the social and emotional development of children to develop an evaluation design that would do justice to Head Start. A test designed on middle class standards of emotional and social development would not measure the impact Head Start has had on an economically disadvantaged child and his family.

### Conclusion

I recommend that OCD and Rand give serious consideration to developing a simplified evaluation design that would give fair, accurate, and non-biased data on the effectiveness of the Head Start program toward improving the social competence of a Head Start child and his family. If social competence is the overall goal of Head Start, I also recommend that serious consideration be given to evaluating and

educating the school systems on the overall goal of Head Start. This is extremely important if the Head Start intervention is to have lasting effects on the child and his family.

I further feel that measures and outcomes should be contingent upon OCD's and Rand's position on developing a relevant evaluation design that could measure the effectiveness of Head Start.

Appendix D  
CONTRIBUTIONS MADE BY PANEL OF SPANISH-  
SURNAMED PROFESSIONALS

This appendix is composed of reaction papers (to Rand's Interim Report) prepared by participants in a panel convened by the Office of Child Development, Department of Health, Education, and Welfare, and The Rand Corporation in January, 1974. The reaction papers were addressed and returned to Dr. Ramón García at the Office of Child Development.

**SOUTHWEST EDUCATIONAL DEVELOPMENT LABORATORY**  
**211 EAST SEVENTH STREET, AUSTIN, TEXAS 78701 512/476-6861**

**February 7, 1974**

**President**  
**NORMAN FRANCIS, President**  
Xavier University  
New Orleans, Louisiana

**Vice-President**  
**WAYNE H. HOLTZMAN, President**  
Hogg Foundation for Mental Health  
Austin, Texas

**Secretary-Treasurer**  
**J. D. LANCASTER, JR., Superintendent**  
Ouachita Parish Schools  
Monroe, Louisiana

**Directors**  
**JAMAR W. ADCOCK, Chairman of Board**  
First Fidelity Mortgage Company  
Monroe, Louisiana

**JACK L. DAVIDSON, Superintendent**  
Austin ISD  
Austin, Texas

**RAFAEL H. FLORES, Attorney**  
McAllen, Texas

**GORDON FLORY**  
Executive Vice-President  
Louisiana AFL-CIO  
Baton Rouge, Louisiana

**RAYMOND FLOYD**  
Professor of Education  
Southern University  
Baton Rouge, Louisiana

**R. A. GARDEA, M.D.**  
El Paso, Texas

**JOHN E. GRAY**  
President  
Lamar University  
Beaumont, Texas

**ABNER HAYNES, President**  
Southwest Minority Corporation  
Dallas, Texas

**ALMA B. HENRY, Supervisor**  
Staff Development  
Orleans Parish Schools  
New Orleans, Louisiana

**SISTER CAROLEEN HENSEN**  
Superintendent of Schools  
Diocese of Dallas  
Dallas, Texas

**WALTER C. MCGEE, III,**  
Consultant  
Houston, Texas

**PAUL MOORE, Superintendent**  
Caddo Parish Schools  
Lake Charles, Louisiana

**ARNULFO L. OLIVEIRA, President**  
Texas Southwest College  
Brownsville, Texas

**W. M. PATTERSON, President**  
Ile de Southern Company  
Itholder, Louisiana

**LEONEL PELLEGRIN, Director Emeritus**  
Division of Continuing Education  
Louisiana State University  
Baton Rouge, Louisiana

**DANIEL SAUCEDO, Director**  
Imman Christian Center  
San Antonio, Texas

**EDWARD J. STEINEL, Executive Director**  
Public Affairs Research Council  
Baton Rouge, Louisiana

**DANA WILLIAMS, Superintendent**  
Cypress Christi ISD  
Corpus Christi, Texas

**Executive Director**  
**JAMES H. PERRY**

**Dr. Ramon Garcia, Research Associate**  
**Office of Child Development**  
**P. O. Box 1182**  
**Washington, D. C. 20013**

**Dear Ramon:**

I want to thank you very much for the grand time which you engineered for us during the OCD-RAND meeting on its proposed design to evaluate social competency in Head Start children.

I wish to summarize my observations regarding that design in this letter.

First of all, I would like to urge you and the OCD staff in charge to ask Rand to submit another interim report reflecting the concerns of the Black and Spanish-speaking groups which met with them to discuss certain inadequacies of their proposed design.

Regarding some of the variables they propose to study, I suggest that they also look at ethnic identity, teamwork versus competition (in appropriate settings which would require teamwork or competition), and the sharing of achievement with others.

Under the category of "Mental and Physical Health," I suggest that they also look at ability to concentrate on self- or group-initiated activities, sense of responsibility toward others, and a sense of order.

Under their general rubric of measuring verbal and non-verbal interactions, I suggest that they acquire the services of someone who could help them understand the possible larger set of meanings to what some observers might perceive to be "the same word." It is possible that other words in the utterance or that certain types of gesticulation accompanying an utterance change its meaning and thereby expand the child's transactional vocabulary.

February 7, 1974  
Dr. Ramon Garcia  
Page 2

A related topic has to do with the measurement of language dominance as a covariant with other independent variables. Another independent variable may have to do with the percentage of ethnic minority students in a program, and the non-dominant ethnic groups that they represent. Also, the design must be sensitive to certain special classifications within ethnic groups which might effect program outcomes, e.g., migrants.

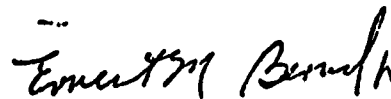
Of special interest to me is the provisions which they will make in the design for the development of cross-ethnically valid instruments. You will recall that I gave them some references to look at this in this regard, but I still believe that some creative thinking on their part will be necessary to guarantee that any future evaluator of the Head Start Program does in fact have some procedures to guide him in delivering culturally appropriate instruments.

This brings up the whole issue of a minority child's social competence in two cultural settings. One of the questions which I feel must be asked by the ultimate evaluation of Head Start is whether certain school-related skills are learned at the expense of a child's social relationships, and, by implication, of a child's mental health--the whole business of unanticipated and undesirable consequences. The question may be extended further by asking whether there are equivalent specific behaviors which a youngster may evidence in order to indicate, say, certain levels of physical development, or whether the indicators will conform to the kinds of activities and skills that middle class dominant-group children perform. Child development specialists knowledgeable about the several cultures may be able to assist in this planning.

Finally, any predictive validity which Head Start measures may have for treatment and control group children entering traditional programs may be different if these children enter highly innovative programs in the first grade or if they enter ostensibly ethnically-linguistically compatible programs, such as bilingual education.

I look forward to hearing from you about this matter in the near future. Please let me know if I may be of further service.

Very truly yours;



Ernest M. Bernal, Jr., Ph.D.  
Director  
Bilingual Early Elementary Program

EMBJ/mac



# The Child Development Associate Consortium

February 7, 1974

Dr. Ramón García  
Research Associate  
Office of Child Development  
P.O. Box 1182  
Washington, D.C. 20013

Dear Dr. García:

Enclosed please find a copy of my "reaction paper" to the Rand Corporation study on social competence.

Let me commend you for your active endeavor to provide OCD with a Chicano perspective. It was indeed a pleasure to be part of such an outstanding working group.

If I can be of any further assistance, please do not hesitate to call me.

Hasta luego,

Josue Cruz, Jr.

Assistant Director - Assessment

JC/dp  
enclosure

REACTION PAPER

Submitted by:

Josue Cruz, Jr.  
7315 Wisconsin Avenue  
Suite 601-East  
Washington, D.C. 20014  
(301) 652-7144

TABLE OF CONTENTS

- I. Introduction
- II. Implications for the Spanish-speaking Child
- III. Organizing Conceptual Framework
- IV. Language Development
- V. Recommendations for Further Study



## I. Introduction

On January 22-24, 1974, Dr. Ramón García, Research Associate at OCD, convened a group of educators/researchers to review a proposal and interim report submitted to OCD by the Rand Corporation. As a result of those sessions with Dr. García and representatives from Rand, I have decided to highlight some of my concerns and observations. I trust that the following pages will help in the development of an "appropriate" and "acceptable" evaluation design. It should be emphasized that the observations at hand are very basic and should be dealt with in much more depth than appear in this document.

## II. Implications for the Spanish-speaking Child

There is absolutely no argument that the study by Rand can have enormous implications for the Spanish-speaking child. The fact that fifteen percent of all children in Head Start are Spanish-speaking clearly indicates the importance of such a study.

Referring to the interim report, it is obvious that the five areas of outcomes affect all children regardless of background. But it should be kept in mind that certain groups of children demonstrate the outcomes in many different ways. If a person accepts the contents of the interim report as a frame of reference for an national evaluation design, there would be great harm to the Spanish-speaking child. The report appears to be very white and middle-class, the overwhelming majority of Spanish-speaking Head Starters are brown, lower-class and speak a language other than English.

One can draw many implications of such a study for the Spanish-speaking, but unfortunately the majority of them are negative and detrimental to the young child. In order to design an evaluation process sensitive to the Spanish-speaker, the following questions must be answered:

- A. What are the unique learning experiences of children in diverse cultures which can be used as a base and means for further development?
- B. What differences between home and school make a difference in linguistic and cognitive development?
- C. What constitutes a positive self-concept for children in each cultural community and how can it be fostered?
- D. Which adult behaviors interfere with the normal developmental pattern in children from different cultures, and which enhance it?
- E. How can children learn to express themselves in a second language and culture without experiencing interference between the two?
- F. To what extent do nutritional and other factors arising out of socioeconomic differences affect learning potential (and values)?
- G. How can negative attitudes and expectations be changed?

Once these questions have been answered, the next step should be the development of a matrix to adequately document those social competencies characteristic of the Spanish-speaking child.

### III. Organizing Conceptual Framework

An important element missing from the interim report is a conceptual framework. The following matrix can serve as a point of departure in utilizing the various components of a conceptual framework.

Cause Outcomes	Culture	Setting	Learning Styles
Health & Nutrition			
Motor Per- ceptual Devel.			
Language Devel.			
Cognitive Devel.			
Socio- Emotional Devel.			

The matrix is an easy tool to utilize and will certainly capture those outcomes characteristic of the Spanish-speaking child.

### IV. Language Development\*

Everyone who lives in the Southwestern United States "understands" Mexican-Americans, their values, their problems, and their life-style. To prepare teachers of Mexican-American children (in case the teachers come from another part of the country or lack assurance in their classroom practice), the educational literature and local curriculum guides provide handy lists of these cultural traits. Mexican-Americans are reported as typically:

- a. "Passive. They accept their poor lot in life, saying, "Que será, será."
- b. Non-competitive. They lead a peaceful rural existence and do not care to join the urban rat race. Nor is much attention paid to such competitive aspects of school as test scores and grades.

\*Information for this section is taken primarily from "The Mexican-American Preschool Child: A Report on Current Research," by Rosario C. Gingräs.

- c. Present-oriented. They work to satisfy present needs, and not for future goals. (This explains the lack of importance the family places on their children's education.)

Many Mexican-American children are considered "alingual" as well - without language. They speak only a mixture of Spanish and English, really neither one, or they don't talk at all. This is blamed on their noisy, crowded home environment and the number of children in each family, which prevents the mother from talking much to any one of them.

In fact, almost no generalizations about Mexican-Americans can be substantiated by objective research if one does not begin with the invalid assumption that the "Mexican-American culture" is a monolithic whole. There are important regional, social class, and rural/urban differences in the population which are seldom taken into account when data are reported.

Little is really known about the values Mexican-American children learn by being members of that ethnic group. The passive stereotype is commonly applied by the dominant group in a society to minorities, and it may reflect a coping style developed by historically oppressed people in this country to avoid calling attention to themselves or "getting into trouble." It is interesting that the stereotype is being maintained even while such formerly "passive" minorities as Blacks, Chicanos, and Indians are rapidly changing to a much more active coping style.

This contradiction was dramatized several years ago when a well-known Mexican psychologist reported on his cross-cultural studies during a meeting at the National University in Mexico City. At the very time he was presenting his statistically-impressive evidence that Mexicans are passive, the University was just beginning to recover from a full-scale student riot. His conclusions were drawn from the response to such multiple-choice questions as, "What would you do in case of an earthquake?" Texas Anglo students included in the study were judged "active" for responding that they would run outside, while Mexicans were judged "passive" for responding that they would stay inside. These responses only prove that Mexicans know more about earthquakes than do Texans. Californians, too, stay inside in doorways during an earthquake if they are in an area of tall buildings, and it usually takes only one such experience for their children to learn such "passivity." (Such fallacious interpretations of data clearly show the general need in research to be sensitive to cultural bias.)

A study of the cooperative vs. competitive behavior of Anglo, Black, Mexican-American, and Mexican elementary school children (Madsen and Shapira 1970) shows the Anglos and Blacks most competitive with Mexican-Americans somewhat less, but still much more so than the Mexicans. This may well be an urban/rural difference instead of an ethnic one, however, since the Mexican group was rural and only about 20% of the Mexican-Americans in the Southwest still live in rural areas. It may also be a social-class difference, since Wasserman (1971) reports more cooperative behavior among "blue-collar children" - whether Mexican-American, Black, or Anglo. Another study by Kagan and Madsen (1971) included four and five year old children and showed no differences at all at that age. Only 3% of the moves of each group in the test rates "competitive," and no group behavioral differences appeared along this dimension until age seven to nine. Yet another study (Del Campo 1970) finds that Mexican-American children score higher on competitive values than do Anglos.

Concern for present needs rather than future gratification is a well-documented characteristic of those who live in poverty. There is absolutely no basis for the stereotype which attributes this to Mexican-Americans as an ethnic group. As for parental interest in education, Mexican-American families place about the same emphasis on education as any other families (Anderson and Johnson 1971). Educators must look elsewhere to explain the high degree of academic failure these children experience.

Claims that Mexican-American children are "alingual" are based on inappropriate testing techniques and misunderstanding about the nature of language and linguistic diversity.

Most Mexican-American children already have two well-developed language systems before they enter school, although they may speak "nonstandard" dialects of one or both. It is quite natural for them to switch from one language to the other, as do adults in bilingual communities, although they should also learn to keep the two codes separate on more formal occasions as they mature linguistically. What is sometimes called "Tex-Mex" is a regional variant of Spanish with some English borrowings in the lexicon. Words like troca for 'truck' also occur commonly in the Spanish of Northern Mexico.

Even older children's Spanish is often deprecated by educators who do not understand the nature of language. An educator in Texas writes:

He speaks Spanish with his playmates. But it is an impoverished Spanish, a language which has been culturally "beheaded" by its forced separation from its own literary heritage.

Another, from Arizona, says:

The fact that the pupil's home language is a colloquial Spanish may be only one additional handicap, no more important than other cultural handicaps.

Conclusions that Mexican-American children's English is stronger than their Spanish (e.g. Cornejo 1973 and Swanson and DeBlassie 1971) may also be based on linguistic naiveté. Those who always speak one language at home and always speak the other in a different domain (like school or work) learn the vocabulary for each domain only in the relevant language. A child may know only the Spanish terms for furniture or cooking utensils found at home, for instance, and only English for such uniquely scholastic objects as chalkboard and filmstrip projector, or terms in subjects like geography and science which he might never discuss at home. Even bilingual teachers who were educated themselves in monolingual English schools have experienced considerable initial difficulty teaching these subjects in Spanish.

Intelligence and achievement tests in Spanish (particularly when normed in Puerto Rico or Mexico) are often just as inappropriate for these children as those in English, and are just as unreliable. It is little wonder that so many studies find Mexican-American children have a lower IQ than Anglo children.

Although there are several reliable descriptions of the language of Mexican-American children (Lastra 1969, Carrow 1971, González 1970, 1973), there are serious needs for further research, including:

- a. Studies of regional and social variation in adult Spanish-speaking communities in the United States. (Child language needs to be described in terms of the adult speech around him, and not a different norm.)
- b. Studies of language development in the same region by children from different socio-economic levels.
- c. Studies of code-switching phenomena, by adults and by children, and in different contexts.
- d. Studies of the acquisition of social rules governing language use.
- e. Studies of second language acquisition controlled for age, socio-economic level, and learning context.

Denying the stereotypes that appear in the educational literature does not mean denying all differences. The "average" Mexican-American family does differ from the "average" Anglo family in size, occupational level, and economic status; a larger percentage of the Mexican-American population belongs to the Catholic church; a larger percentage maintains bilingual competencies than any other ethnic group; and a disproportionate number of Mexican-American children do fail in school.

Those who fail are most likely to be different from the mainstream Anglo norms in most of these respects, by definition the "unacculturated". A primary goal of early childhood programs has often been to try to eliminate these differences, to change the children and/or their families to fit the educational system they will enter. An alternate possibility, at least theoretically proposed by advocates of bilingual/bicultural programs, is to accept and build upon individual and social differences, to change the educational system to fit linguistically and culturally diverse children.



The following literature survey is arranged by major topics. Within each topic, the documents surveyed will be presented in chronological order.

## **I.- INTELLIGENCE MEASUREMENT**

Most of the literature concerning Mexican-Americans from 1950 to 1961 is summarized in Darcy (1963). In essence, Darcy states:

studies of Spanish-English bilinguals ... have indicated that bilingual subjects received lower scores on verbal tests of intelligence than on nonverbal intelligence tests. In some studies the translation of the directions of intelligence tests from English to Spanish did not improve the test scores of bilingual subjects, while in another investigation, in which the subjects had received several years of formal education in English and had a poor knowledge of Spanish, mean scores on the English version of the intelligence test were significantly higher than were mean scores on the Spanish translation of the test (Keston & Jiminez 1954).

Most of the documentation after 1963 involving intelligence measurement and Mexican-Americans seems to be more concerned with testing the testing-instrument than in describing the subjects. That is, since Mexican-Americans consistently do worse as a group than the Anglo-American group, something must be wrong with the test.

1. T. Quijano, "A cross-culture study of sex differences among first-graders on a verbal test," 1968.

This document describes an experiment on thirty Mexican-American and Mexican six-year-olds. The goal was to see if there were any significant differences in the results of a test of verbal ability between sexes. The instrument used was the Van Alstyne Picture Vocabulary Test. The test was translated into standard Spanish and what the experimenter called "Tex-Mex". The results of the experiment were that there were no significant differences between sexes on this test of verbal ability.

This experiment is interesting in that recognition is made of the fact that the Spanish spoken in Texas may be different from that spoken in Mexico. However, what the experimenter labels Tex-Mex is simply standard Spanish with a few English borrowings thrown in (e.g., mecha (English 'match') instead of standard Spanish 'fósforo'). The "Tex-Mex" version of the test was given to children in Laredo, Texas, a border city heavily influenced by Mexican culture. Although the experimenter claims that there are no verbal testing instruments available in Spanish, it is unlikely that the "Tex-Mex" version could be used outside

of the immediate border area (e.g., Los Angeles or Albuquerque). The experiment is interesting only in that this is the first attempt to provide a measuring instrument for Mexican-Americans. However, it does not seem that it would work out successfully. The results of the experiment may be interesting if true. The experiment does not give a convincing argument to conclude that there is no difference in verbal ability between sexes for Mexican-Americans at age six. The sample base is too small (and not specified).

2. F. Karabinus and M. Hurt, Jr. "The Van Alstyne Picture Vocabulary Test used with six-year-old Mexican-American children," 1969.

Two groups of six-year-old Mexican-American children were tested in 1965 and 1966 (N=535) using the Van Alstyne Picture Vocabulary Test to measure intelligence. No demographic information is given on the subjects other than the statement that they were disadvantaged. The results indicate that the Van Alstyne Picture Vocabulary Test is reliable and valid for the measurement of mental ability of culturally disadvantaged Mexican-American six-year-olds. Comparison of the results with data in the test manual (of the Van Alstyne Picture Vocabulary Test) based on 93 six-year-old children selected from the general population showed all reliability coefficients calculated with the scores of the subjects higher than .71 (Spearman-Brown) found in the general norming population. This document tells little about Mexican-American children, but does show that at least one measuring instrument exists that may be free from cultural bias.

3. T. Christianson and G. Livermore, "A comparison of Anglo-American and Spanish-American children on the WISC," 1970.

The purpose of this study was to compare the performance of lower and middle class Anglo-American with lower and middle class Spanish-American subjects on the Wechsler Intelligence Scale for Children (WISC). The subjects for the study were 92 Anglo-American and Spanish-American children 13 to 14 years of age. The subjects were classified, on the basis of social class and ethnic origin, into four groups of 23 children each. The data show middle class children in both ethnic groups scored significantly higher than lower class children on each of the WISC measures examined in the study. On those measures where ethnic origin was a factor, Anglo-Americans scored significantly higher than Spanish-Americans. The results showed that:

- general intelligence and the development of verbal abilities, including the ability to use acquired verbal skills in new situations, are related to ethnic origin and social class. Nonverbal abilities, perceptual organization ability, and the ability to concentrate on a task were found to relate only to S's membership in a particular social class (pp. 12-13).

THE RELATIONSHIP OF SOCIAL CLASS AND ETHNIC ORIGIN TO SCORES ON THE WISC  
(p. 12)

	Group means		Spanish-Am. middle cl.	Anglo-Am. middle cl.	F values ethnic orig.	soci clas
	Spanish-Am. lower class	Angl-Am. lower class				
WISC measures						
Full scale IQ	91	99	111	116	12.93*	95.8
Verbal scale IQ	89	95	111	120	12.66*	82.8
Performance scale						
IQ	96	102	108	109	3.60 <sup>a</sup>	28.8
Verbal comprehension	8.1	10.0	12.2	13.9	12.00*	30.9
Freedom from dis- tractibility	8.8	8.6	11.7	11.0	.88 <sup>a</sup>	32.7
Perceptual organization	9.5	10.3	11.4	12.1	3.02 <sup>a</sup>	16.4

Note: <sup>a</sup>= not significant  
\*Significant at the .01 level.

The significance of this study lies in the fact that it is one of the few that observes Mexican-Americans in terms of different social classes (although the subjects are labeled Spanish-Americans, their location is not specified, and the sociological parameters used to classify the children into social classes are not discussed). Although the age group is 13 and 14 years of age, it does give a hint as to what might be expected if other age groups were tested. Although Mexican-Americans still perform poorer as a group than Anglo-Americans on the WISC (possibly the WISC is measuring linguistic ability in English), it is significant that the middle class Mexican-American group performs better than the lower class Anglo-American group.

4. E. Swanson and R. DeBlassie, "Interpreter effects on the WISC performance of first grade Mexican-American children," 1971.

The purpose of this study was to determine the effect of the use of an interpreter in the administration of an individual intelligence test on the performance of a group of Mexican-American bilingual children. Forty one first grade children between the ages of 6 years 8 months and 7 years 11 months from two elementary schools in a rural school system in central New Mexico were selected as subjects. The California Test of Mental Maturity was administered to all subjects who were then ranked according to Total IQ score results. Alternate subjects were then assigned to one of two groups--an experimental group (N=21) or a control group (N=20). The Experimental Group was administered the WISC by one of the researchers with the use of an interpreter. The children were encouraged to answer in Spanish if they wished. The control group was administered the WISC, but entirely in English.



The results are as follow (p.174):

Means, Standard Deviations, and t's contrasting the Experimental and Control Groups on the WISC

WISC IQ'S	Experimental (N=19)		Control (N=18)		t
	M	SD	M	SD	
Verbal	88.95	14.04	91.94	19.00	.85
Performance	101.42	15.63	103.44	15.66	.39
Full Scale	94.47	14.66	96.94	17.63	.46

The results show that the presence or absence of an interpreter does not result in statistically significant differences. Swanson and DeBlassie note that these results are in direct conflict with the research reported by Mycue (1968). Mycue had reported that a group of Mexican-American children had performed significantly better on the Language Facility Test when it was administered in both Spanish and English than when it was administered in English only (as reported in Swanson and DeBlassie, p. 174). The experimenters, however, urge caution in interpreting the results since the sample base is very small. They advise that further testing be undertaken.

It would appear that competence in English is higher than competence in Spanish for these subjects. That is, even if an interpreter is used, the level of competence is so much higher in English that the use of an interpreter would not allow the subject to significantly increase his level of performance on the WISC.

This agrees with the conclusions reached by Keston and Jiminez (1954). In administering an English and a Spanish version of the Stanford-Binet Intelligence Test (Forms L and M), they found that the subjects (50 Mexican-American fourth graders in Albuquerque) performed more poorly on the Spanish version than on the English one (the mean IQ on the Form M English version was 86.0, while the mean IQ on the Spanish Form L was 71.8). Keston and Jiminez relate this difference to the fact that the level of development in the English language used by the children tested was higher than that in the Spanish language they used. These children received their formal education in English, and since the Stanford-Binet reflects the educational achievement of children, it could be expected that higher scores would be obtained in the language which was more highly developed in formal aspects. It was the examiners' impression that these children had speech habits of preschool children in Spanish conversation. Hence it was suggested that the development of the Spanish language was brought practically to a standstill when the child entered school and began formal education in English. Thus Keston and Jiminez provide a hint as to why an interpreter would not result in any improvement in performance as noted by Swanson and DeBlassie.

5. T. Hickey, "Bilingualism and the measurement of intelligence and verbal learning ability," 1972.

This experimental study attempts to show that one widely used instrument is ineffectual for measuring bilinguals. The instrument is the Peabody Picture Vocabulary Test (PPVT). The results of the experiment show that the subjects (160 monolingual Anglo-Americans and 160 bilingual Mexican-Americans, age 4) that are bilingual encounter great difficulty in correctly identifying verbal noun concepts on the PPVT. Structural and idiomatic differences between English and Spanish were thought to be the source of the difficulty. This is significant in that the PPVT is constructed so as to preclude the conclusion that all preschoolers, bilingual or monolingual, might encounter a great deal of difficulty with verbal noun concepts. The author notes that one particular sequence has a picture of a waterfall and another picture shows a child falling off his skates. It is not clear (from the article) just what responses are expected, except that English is supposed to allow the same type of answer for both pictures, while Spanish requires two different responses. The author is not very specific as to why this should be the case, but he feels that semantic interference phenomena are at play and may result in an error for the bilingual child. No examples of typical responses are given, and no conclusion is reached. The author ends the summary by stating that more studies are necessary to determine just what the differences are that may exist between Spanish and English at the structural level. The author seems to be unaware that a rather large body of information already exists on this topic. The significance of this study appears to be that the PPVT may well be biased against the bilingual and perhaps should not be used with Mexican-Americans. Unfortunately the author does not seem to be familiar with linguistics, and consequently his findings cannot be re-interpreted to provide a more conclusive result.

## II: LINGUISTIC DESCRIPTIONS

There are few descriptions of Mexican-American child language. The few studies that do exist appear to suffer from three general methodological limitations. These are: (1) the child's speech is described in the absence of a description of the speech spoken by adults that may be around him; (2) incipient bilinguals are not distinguished from functional bilinguals; (3) the focus is on the speech of bilinguals to the almost total exclusion of monolingual speech.

The first limitation results in difficulty in distinguishing developmental phenomena from dialectal features. For example, Lastra (1969) observes that East Los Angeles children frequently (usually?) have a labiodental fricative where standard Spanish has a bilabial fricative ([v] in place of [β]). In the absence of a description of adult speech in East Los Angeles, it would seem that [v] might be some type of developmental phenomenon. However, as

Phillips (1972) points out, the substitution of [v] for [ß] is fairly common among adult speakers in Los Angeles. Consequently, this particular contrast with standard Spanish is not an instance of developmental phenomena. Lastra points out that the children observed in her study also seem to have glottal stops between two consecutive vowels [mi'ermano]. Since it probably is not the case that adult speakers of Spanish (in East Los Angeles) usually place glottal stops between vowels, this is very likely to be a developmental characteristic. Without data on adult speech, there is no way to determine this.

The second limitation results in confusing what a child does when he is learning a second language with what he does when he has already internalized two sets of grammatical rules. Learning interferences may not be the same as bilingual code interferences. It would be unlikely that a child learning English would say in Spanish "mi tío's casa" (my uncle's house); a functional bilingual might very well say this as an instance of code-interference (as reported by Lastra).

The third limitation results in trying to describe speech that is most variable of all. A functional bilingual may produce instances of code-interference for any number of reasons--many of a nonlinguistic nature. This concern with the most difficult of speech to describe (because of the inherent instability--the bilingual can mix his codes at will) has resulted in no descriptions of the English spoken by monolingual Mexican-American children. It may be the case that monolingual English-speaking Mexican-American children learn a kind of Hispanified English (as a result of diachronic processes), but at this time one can only guess what kind of English it might be. Killian (1971) has observed that monolingual Mexican-American children do not perform as well on intelligence tests as Anglo-American children (at age 6), but appear to perform better than bilingual Mexican-American children. It may well be the case that the English spoken by the monolingual Mexican-American children is very different from that spoken by the Anglo-American children. Unfortunately no information exists concerning this problem.

Descriptions of how a Spanish-speaking child learns his language (along developmental lines) will be crucial information for describing how a Mexican-American child develops linguistically (assuming that he is to be bilingual).

6. Y. Lastra, "El hablar y la educación de niños de origen mexicano en Los Angeles," 1969.

Lastra investigated the speech of 65 children in East Los Angeles, ages ranging from 5 to 9. The Spanish and the English of these children was observed. 20% of the children were first generation (born in Mexico); 54% were second generation (parent or parents born in Mexico); 25% were third generation (grandparent(s) born in Mexico). Six children spoke what the interviewer considered standard Spanish. These children were

all second generation. Standard speech did not appear to be related to the occupation of the father.

Lastra observes that the majority of the children speak a dialect of Spanish similar to that of lower class speakers of Mexico, although sprinkled with interferences from English. All understand the standard dialect although the children lack vocabulary that would ordinarily be learned in school (mathematics, geography, etc.).

Most of the features that Lastra describes are simply differences between local Los Angeles Spanish and standard Spanish. Among those features which may be developmental in nature are the following:

**Phonological:** Use of a retroflex /r/ before a consonant:  
[kárne], [enférmo].  
Multiple R overly trilled (hypercorrection?).  
The multiple R becomes a tap: [aríba] instead of [arriba].  
Voicing of Sp. /s/ between vowels: José [hozé].  
Glottal stop between two consecutive vowels and in initial position before a vowel: [mi'ermano], ['ermano].  
/g/ remains a stop between vowels: pega [pega],  
amiga [amíga].

**Syntactic:** Absence of usted (exclusive use of tu).  
The use of the article before proper names: la Cecilia,  
el James.  
The use of the form mi as direct object: mi began,  
mi lo compró ["They hit me", "He bought it from me"].

Lastra also describes the English used by these children. Among the things she notes are the following:

**Phonological:** Intonation similar to that of Spanish:  
I want to be<sup>2</sup> a football<sup>2</sup> player<sup>1</sup> (instead of 231).  
Fricative [oßer] "over" [ay8li] "I live".  
Centralization of the lower front vowel: [lamp] "lamp".

**Syntactic:** Mislocated adverbs of time: "Sometimes at night we play games."  
Repetition of the subject: "My mother, she doesn't have a job."  
Double negatives: "Mrs. E is not teaching no more."  
Lack of agreement: "Does Bertha and Sandra play with you at home?"  
Past instead of infinitive: "I used to threw the ball."  
"I haven't gave them a name yet"  
Confusion of the gerundive with the infinitive:  
"I like to doing is math."

Although the data are not presented systematically, nor is any information given about what is being said at what age (the difference between 5 and 9 would seem to be quite great), this study is significant since it is the only study available on the speech of young Mexican-American children in Los Angeles.

7. D. Natalicio and F. Williams, "Repetition as an oral language assessment technique," 1971.

The purpose of this study is to assess the degree to which sentence imitations by Black and Mexican-American children (K-2) could be used as a basis for language evaluation. A panel of "experts" was commissioned to assess the sentence imitations. Since no clear criteria were established for the panel, it is very difficult to judge the efficacy of the project. However, incidental to the purposes of the project, the authors do give some hints as to the speech of the children. Ten Mexican-American children (6-8 years of age) were selected from 750 subjects in San Antonio. No further demographic information is given. The subjects were to repeat a series of sentences both in Spanish and English. Following is a list of the types of difficulties the children had:

English:

- Difficulty with the prepositions on, in, and at.  
/s/ and /c/ were 'confused'.
- Devoicing of /z/ to [s].
- Third person ending /-z/, /-s/, /-iz/ deleted on repetition.
- Difficulties with /r/ and /l/ (liquids; also in Spanish).
- No aspiration of initial voiceless stops.
- Schwa replaced by [a].
- Substitution of you, your for his, her, their.
- Substitution of one for a(n) (David has one brush for his hair).

Spanish:

- 'Misuse' of reflexive se: 'Los ninos [se] acuestan',  
'David se puede abotonarse la camisa',  
'David se puede abotonar la camisa' (instead of 'David puede abotonarse la camisa').
- /l/ substitutes for /r/ in final position: [tlabaxal] 'trabajar'.
- Reduction of /ye/ and /yo/ diphthongs: tene (for 'tiene').  
meto (for 'metio').
- /-y-/ substituted for /-d-/ and /-r-/: gloya (for 'gloria').  
ayuya (for 'avuda').  
yentes ('dientes').

Deletion of redundant dative object:

Mama ayuda a Gloria. (Mama le ayuda a Gloria.)

El jabón se metió en los ojos.

(St. Sp.: El jabon se le metio en los ojos.)

Preposition a replaced by pa(rá): Ellos van pa la escuela hoy.

Preposition en replaced by a: Se le metió a los ojos.

The authors certainly cannot be faulted for presenting only tid-bits of data, since this was not their intent.



8. R. Serrano, "The language of the four-year-old Chicano," 1971.

The author claims that "little is known about Chicano four-year old language." His article (based on samples of 14 subjects) does not add much to our knowledge. The article is highly inconclusive, but does give some additional bits of data. There is no demographic information on the subjects (not even their locale). It is not clear whether the subjects are still learning English or have already acquired a degree of competence. Most of the samples would be expected from someone learning English (/e/ and /æ/ merger, etc.). The most interesting point Serrano makes is that his subjects do not distinguish between tell me and ask me (this was first noted by Carol Chomsky).

9. E. Carrow, "Comprehension of English and Spanish by preschool Mexican-American children," 1971.

The purposes of this study were:

- (1) to compare the comprehension of English with that of Spanish in a group of preschool Mexican-American children;
- (2) to analyze the developmental sequences of the two languages in these children;
- (3) to compare these developmental sequences of both languages in the children under study with the performance of a group of English-speaking children reported earlier.

Ninety-nine children with Mexican-American surnames, ages 3-10 to 6-9 were tested in Houston. Each child was administered the Auditory Test for Language Comprehension---an instrument that allows the assessment of oral language comprehension without requiring language expression from the child.

The interesting part of the study is the data showing when certain linguistic features are first understood (at the 60% level, i.e., 60% or more of the children in a given age group comprehended the test item). Among the most significant findings (significant because of syntactic problems; most of the test items involve lexical items) are the following:

At no age level up to 6-0 years did as many as 60% of the children understand in English or Spanish the pronouns 'he' ('el'), 'her' ('de ella') and 'his' ('de el') as contrasted with 'she' ('ella'), 'his' ('de el'), and 'her' ('de ella'), respectively, although these contrasts were comprehended by 60% of the control children [Anglo-American] at 4-0.

Sixty percent of the experimental group at all age levels the preposition "on" ("sobre"). There was a year's delay in comprehension of "under" ("de bajo de") and "in" ("en"), as compared with controls. Two prepositions that were considerably delayed in English, "by" ("al lado de") and "in front of" ("en frente de") were equally delayed in Spanish.

The controls understood "is not" and "isn't" at 3-0 and 3-6 respectively, while the children in this study did not comprehend these negatives until 6-0 in English... The only tense contrast the experimental group was able to comprehend in either language was the present progressive. The control group understood the past tense at 4-0 years, the future at 4-6, and the past progressive at 6-0.

The demonstratives "that" ("aquel") and "these" ("estos") were comprehended at all ages in English, but not until 6-0 in Spanish. The interrogative "who" ("quien") was not understood in either language by 60% of the children, although the control group understood it at 3-0.

Structures of predication and modification involving complex syntactic relationships (complex sentence with independent clause and dependent adjectival clause, complex imperative sentence with conditional clause, etc.) were comprehended at the same age as the control children [age not given]. The most difficult structural contrasts were those of direct-indirect object and active-passive voice.

The items which ranked within the last ten in both languages, that is, those understood by fewest of all the children, were the lexical items "few" ("pocos"), "alike" ("igual"), "different" ("diferente"), ... and "pair" ("par") and the following structural items: "Neither the boy nor the girl is jumping" ("Ni el muchacho ni la muchacha está saltando"), (negative); "Who is by the table?" ("Quién está al lado de la mesa?") and "The man is hit by the boy" ("El hombre es golpeado por el muchacho"). All the previous items except for "pair" (4-0) and "who" (3-0) were passed at 6-0 and 7-0 by the controls.

The implications of the study, as seen by the author, are (1) the Mexican-American children are a very heterogeneous group; (2) among preschool children from a low socio-economic level in Houston, the greater proportion understand English better than Spanish. It is probable that this fact would be even more pronounced among middle and upper class Mexican-American children; (3) in general both languages improve as the children become older.

The study concludes by stating that some of the major problems for these children seem to be in pronominal reference, negation, tense marker comprehension, adjectives, prepositions and pluralization.

This study is probably the most significant done so far on the Mexican-American preschool child. However, some of the language used in the test items is open to discussion. For example, the frequency of passive constructions are not the same for both Spanish and English.





literary sources (e.g., Thorndike's and Rodriguez')? Just whose speech is represented by the baby talk inventory? One child, 5, or \_\_\_\_? If any one child had all the features listed in the baby talk lexicon, would he be understood? There is no demographic information on the children (except for geographical origin). No contrast is made between the speech in the community (adult) with that of the children. Why is the elimination of certain final consonant clusters unique to these children (in English)? Is it not the case that most English-speakers delete /-d/ after a nasal before juncture in South Texas? This is a good example of a mish-mash of developmental phenomena and dialectal characteristics.

### III: VARIA

11. A. Jensen, "Learning abilities in Mexican-American and Anglo-American children," 1961.

Mexican-American and Anglo-American fourth and sixth graders of different IQ levels were compared on a number of learning tasks consisting of immediate recall, serial learning and paired-associates learning of familiar and abstract objects. The results are that on the direct measures of learning ability used in the study, Anglo-American children of low IQ are slow learners as compared with Mexican-Americans of the same IQ. Mexican-Americans of above average IQ do not differ significantly in learning ability from Anglo-Americans of the same IQ. The study suggests that the majority of Mexican-Americans with low IQ's, at least as measured by the California Test of Mental Maturity, are actually quite normal in basic learning ability, though they may be poor in scholastic performance for other reasons. A low IQ in the Anglo-American group, on the other hand, is in most cases a valid indication of poor learning ability.

Jensen advises that most of the low IQ Mexican-Americans, not being basically slow learners, not be placed with Anglo-Americans of low IQs.

12. R.W. Henderson and C.B. Merritt, "Environmental backgrounds of Mexican-American children with different potentials for school success," 1968.

This study is an investigation of the extent to which a wide range of environmental stimuli are differentially present in the backgrounds of Mexican-American children who have, respectively, relatively high and relatively low potential for success in school. One group of 38 six-year-old Mexican-American children in Tucson were tested by the Goodenough-Harris Drawing Test and Van Alstyne Picture Vocabulary Test and had the highest scores on these tests. This group was designated high potential. The low potential group also consisted of Mexican-American six-year-olds (N=42) who had received the lowest scores on the two above mentioned tests. Trained interviewers employed an

interview schedule and a focused interview technique to secure ratings on 33 characteristics defining a set of nine environmental process variables. The mothers of the subjects received the interviews. The data indicate that children in the high potential group come from backgrounds that offer a greater variety of stimulating experiences than is available to most children in the low potential group. The children in the high potential group scored high in the vocabulary test. The low potential children come from an environment with many siblings. This suggests that mothers with many children have less time to interact with any particular child. The author feels that this might explain why so many Mexican-American children come to school with little linguistic ability in either Spanish or English.

13. R.J. Malaragno and G. Newark, "A pilot study to apply evaluation-revision procedures in first grade Mexican-American classrooms," 1968.

This article is concerned with testing a teaching technique called "evaluation-revision strategy." What is interesting, if true, is that the authors isolated a set of seventeen 'concept words' that first grade Mexican-American children appear to have trouble understanding when learning to read. Unfortunately, the authors only identify twelve of these words: top, bottom, alike, different, first, middle, last, under, over, underline, on, and above. The researchers mention that the children (no number) were tested for knowledge of these concepts in Spanish (with no indication as to how they were translated, nor of the results of the test). It is interesting to note that some of these words also occur in the Carrow study and caused problems to the children in that study.

14. Ruth Silverstein, "Risk-taking behavior in preschool children from three ethnic backgrounds," 1969.

This study was undertaken to examine the basis of the unresponsive classroom behavior noted with Mexican-American children. The subjects were 60 Negro children, 79 Mexican-American children, and 25 Anglo-American children 50-62 months in age. No other demographic information is provided. Two hypotheses are proposed:

- H<sub>1</sub>: risk-taking is related to ethnicity, and Mexican-American children take less risks than Negroes or Anglo-Americans;
- H<sub>2</sub>: if risk is related to reward, the Mexican-American child will still take less chances.

The children are given a risk-taking test. The results do not support the hypothesis that Mexican-American four-year-olds exhibit a more cautious risk-taking style than do Negro and Anglo-American children from the same SES group. When candy is rewarded in comparison to beads or praise (on the test), the Mexican-American preschool children exhibit more cautious behavior than the other groups. Since candy is a tangible object (and desirable), this suggests that the motive to avoid failure

seems to be operative for the Mexican-American child (the rationale is amply discussed in the document). The author concludes that the Anglo-American and Negro child is characterized by the problem-solving strategy of "achieving" while the Mexican-American child is characterized by the strategy of "not failing".

This experiment should be replicated with a larger sample base. If it can be shown that Mexican-American preschool children tend to be passive in the classroom because of this strategy of "not failing", then this would go far to explain why these children get "turned off" from school since due to linguistic problems, the scholastic "dice" are loaded against him.

15. M. C. Madsen and A. Shapira. "Cooperative and competitive behavior of urban Afro-American, Anglo-American, Mexican-American and Mexican village children." 1970.

This report discusses four experiments to examine the cooperative or competitive behavior of three U.S. ethnic groups and one Mexican group. The subjects for the first three experiments that tested the three U.S. ethnic groups numbered 144 and were in either second or third grade (ages 7-9). The children were divided into three even groups according to ethnic identity. The measuring instrument was a cooperation board, developed by Madsen. (The board involves pulling strings at each end of a box; cooperation is required of the subjects.) The results indicate that the three U.S. groups responded in a non-adaptive competitive manner over the four trials (in experiment I). The performance of the Mexican-American subjects, although competitive, was consistently less vigorous than the other two U.S. groups. The Mexican Village children were cooperative in their behavior and contrasted sharply with the U.S. groups (which includes the Mexican-American).

This document is interesting in that it shows that the urban Mexican-American child is very different from the rural village Mexican child. This should help in changing some of the stereotype image of the Mexican-American.

16. E. Hopper. "Self-concepts, values and needs of Mexican-American underachievers or (must the Mexican-American child adopt a self-concept that fits the American school?)" 1970.

This somewhat polemical article has a data base of 150 Mexican-American boys aged 9. It is not clear whether the data support the conclusions, but the conclusions are certainly interesting. Her conclusions are:

1. Contrary to stereotyped views, Mexican-American boys even underachieving in school, do not perceive themselves as more negative than their Anglo peers or their better-achieving Mexican-American peers.

2. Contrary to many educators' perceptions, Mexican-American boys, although underachieving in school, value grades and education, and do not consider themselves "dumb in school".
3. Differing again from other studies, this investigation found evidence that Mexican-American boys not only do not have lower occupational aspirations than their Anglo peers, but, in effect, evidenced higher occupational goals for themselves than the other groups with which they were compared.
4. Mexican-American pupils in this study appeared more self-accepting than their Anglo peers or their achieving Mexican-American peers. There were no noticeable discrepancies between their real and their ideal selves and therefore this is taken as another evidence of the lack of low self-esteem of these youngsters.
5. Mexican-American boys are more active than their Anglo peers.
6. Mexican-American boys aged 9-13 do not value "reading for its own sake"--a value featured in the typical American school. They also reject the American cultural value of "leadership".
7. Mexican-American boys do not feel they are as bright as their Anglo peers, and seem to have internalized and accepted the school's view of them with regard to intelligence---even when they are in reality as bright as the other boys.
8. Mexican-American underachievers are most significantly different in the areas compared in this study from Anglo-American boys observed---while Mexican-American achieving boys are more like their Anglo counterparts than their own ethnic peer group.

The reader should be warned that these conclusions are not warranted by the data (or lack thereof) presented in the article.

17. S. Wasserman, "Values of Mexican-American, Negro and Anglo blue-collar and white-collar children," 1971.

This study investigated relationships between four-year-olds' expressed humanitarian and success value preferences and their ethnicity, socio-economic status, and sex. The humanitarian values investigated were helpfulness, cooperation, concern for others, and sharing; the success values were competition, status, expertise-seeking, and completion of task. The sample consisted of 180 children and included equal numbers of 4-year-old Mexican-American, Negro, and Anglo children. Each ethnic group was composed of an equal number of white-collar and blue-collar children (30 each). The instruments consisted of 16 pictures depicting value conflict situations. The interviewer told an accompanying story with each picture. Wasserman reports that the direction of the differences of children's expressed humanitarian values and expressed success values indicate that the Anglo children's scores were higher than those of the

Mexican-American and Negro children. No significant comparison of Anglo and Mexican-American children was shown in gauging humanitarian value, but significant differences were found between scores of Anglo and Mexican-American children in scores for success value complex. The scores of the blue-collar children were higher for cooperation than those of white-collar children.

18. G.H. Naylor, "Learning styles at six years in two ethnic groups in a disadvantaged area," 1971.

The purpose of this study was to investigate learning styles among young children of a disadvantaged area. The subjects were 40 Mexican-American and 40 Anglo-American 6-year-old children in first grade. Tests were selected to measure four learning styles: (1) information demand; (2) impulsivity-reflectivity; (3) field independence-dependence; and (4) originality. No significant differences between sample groups were found on the basis of learning style test performance with the exception of the impulsivity measure where the Anglo group made more errors than did the Mexican-American group.

Naylor points out that differences in the learning behavior of Mexican-American children when compared to Anglo-American children have commonly been attributed to the differential influence of the Mexican-American culture. In this investigation it was expected that the Mexican-American groups would demand more information in decision making, would be more field-dependent, less impulsive, and less original than the Anglo-American groups. The results of these experiments are not entirely conclusive since social class was not taken into consideration. It may be the case that at age 6, middle-class children, be they Mexican-American or Anglo-American, would not be distinguishable on the bases of ethnic origin in respect to the tests for the four learning styles. At any rate, the study indicates that the popular notion of Mexican-American culture may well be more of a myth than is currently being suspected.

19. J. M. Stedman and R. E. McKenzie, "Family factors related to competence in young disadvantaged Mexican-American children," 1971.

The first purpose of this study was to identify high and low competence groups, defined on the basis of linguistic ability and behavioral adjustment, within a population of disadvantaged, preschool, Mexican-American children and then to compare their families on a number of different variables. A second purpose was to investigate the relationship between linguistic ability in both English and Spanish and teacher behavior ratings for the total child sample. The sample consisted of 134 five-year-old Mexican-American children from San Antonio, Texas. Adjustment was measured by The Classroom Behavior Inventory (CBI); language ability by the Tests of Basic Language Competence (TBLC), and the Semantic Differential technique was used to measure the parents' self-concept, roles within the family, and concepts



related to the child's adjustment in school. In brief, low competence parents were found to have a lack of emphasis for formal education (confirming the stereotype); high competence families appear to manifest a semantic and attitudinal structure which is rather "Anglo middle-class" in nature. Although the experiments appear to be inconclusive due to the many parameters employed, it does seem to suggest that even 'disadvantaged' parents differ in their attitudes toward education and Anglo teachers. The heterogeneous nature of the sample population seems to have surprised the investigators. The only value in this experiment seems to be a general warning against over-generalizing about 'disadvantaged' Mexican-American family values toward education.

20. L. Schmidt and J. Gallessich, "Adjustment of Anglo-American and Mexican-American pupils in self-contained and team-teaching classrooms," 1971.

The purpose of this study was to investigate the relationship between school adjustment and three variables: teaching organization, ethnicity and sex. The subjects for this study were taken from predominantly Anglo-American and Mexican-American schools in Texas. First and sixth graders were selected for the study. One hundred and sixty first graders were taken from two team-teaching schools (one of the schools was predominantly Mexican-American) and two self-contained control schools. The six graders also came from four different types of schools (N=383, of which 155 were Mexican-American children). Anxiety of first graders was measured with the Picture Anxiety Test; the anxiety of sixth graders was measured by the Phillips Anxiety Test. The results reported were: (1) Mexican-American first-grade and sixth-grade subjects reported significantly higher anxiety levels than Anglo-American subjects; (2) first-grade subjects in self-contained classrooms appeared more anxious than first-grade subjects in team-teaching settings. Sixth graders were also more anxious in self-contained classrooms; (3) sixth grade females were found to be more anxious than males; (4) the anxiety level of the two ethnic groups did not differ significantly between the two organizations in the first grade, but the anxiety level reported by Mexican-American sixth graders in self-contained classrooms was significantly higher than the anxiety reported by Mexican-American team-taught sixth grade subjects. The main conclusion of the study is that team teaching is not detrimental to the elementary grade children of the study--and may be advantageous for some children.

The interesting results of the experiment is that Mexican-American children have higher anxiety levels than Anglo-American children. Perhaps this is tied to the fear of failure that was discussed by Silverstein (see above).

21. J. Kershner, "Ethnic group differences in children's ability to reproduce direction and orientation," 1972.

The purpose of this study was to see if there are differences in complex visual-spatial ability between Chicano and Anglo children and to see if these differences might be explained partially by the reciprocal polarization of spatial and verbal information processing strategies. Thirty Mexican-American and fifteen Anglo families were selected, each of which had a child enrolled in the local school. Eight bilingual Mexican-American families and eight Anglo families agreed to participate in the experiment. The ages of the children ranged from 6 to 12. The measuring instruments were the Toronto Family Functioning Scale, Warner's Index of Status Characteristics, Slosson Intelligence Test, and Peabody Picture Vocabulary Test.

The results showed that Anglo children are better than Chicano children in matching orally presented words with visual two-dimensional representations of the objects that the words symbolized. Chicano children are better than Anglo children in complex visual-spatial ability. It is not clear why there should be such a distinct difference in information processing strategies. The author claims that one possible reason is that communication between parents and child appeared to be at a minimum (this was observed during the investigators' visit to the homes of the Chicano children). However, it seems if there is little verbal communication between parents and child when the Anglo investigator visits, it may well be due to the presence of the stranger. If the experiments were replicated, and if the same results were obtained, then a clear explanation is necessary for such a distinct difference in strategies.

#### IV. CONCLUSION

As can be seen from the articles reviewed, research on Mexican-American preschool children is limited to two general areas: intelligence measurement and linguistic competence. Most of the remainder of the research does not form a cohesive whole. Although a considerable amount of literature appears to deal with the Mexican-American child, most of it is either incidental to the Mexican-American child, or else it is not too revealing. What is needed is a coordinated effort. Perhaps mainstream America has finally discovered the Mexican-American, but it still has to discover the child.

- A. That a group of Chicano and Puerto Rican consultants remain on call to review all reports and developments affecting the Spanish-speaking.
- B. That Rand respond to the suggestions of the Spanish-speaking consultants as soon as possible with a plan of action prior to beginning any new work.
- C. That Rand subcontract to Chicanos and Puerto Ricans any tasks affecting the Spanish-speaking.
- D. That any contracting of position papers commissioned by Rand on the Spanish-speaking be made with Chicanos or Puerto Ricans.

/dp  
2/6/74



RUDOLPH C. IPOIKE director

CENTER FOR APPLIED LINGUISTICS 1611 North Kent Street, Arlington, Virginia 22209

Telephone (703) 522-4312

Cables CENTAPLING

February 6, 1974

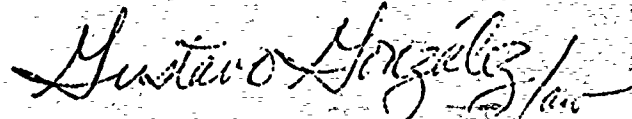
Dr. Ramon Garcia  
Research Associate  
Office of Child Development, OHD  
P.O. Box 1182  
Washington, D.C. 20013

Dear Dr. Garcia:

Enclosed please find a copy of my reaction paper dealing specifically with the language development section.

As I have noted in the last paragraph of the reaction paper, I feel that only in a more extensive treatment can I do justice to the language development section as it relates to the Spanish-speaking population. Please advise me as to how much time will be available for developing an adequate response paper.

Atentamente,



Gustavo Gonzalez, Ph.D., Co-Director  
Bilingual-Multicultural Education

GG:aw

enc

REACTION PAPER

HEAD START LANGUAGE DEVELOPMENT GOALS AND  
THE SPANISH-SPEAKING CHILD

Prepared by

Gustavo Gonzalez, Ph.D.  
Center for Applied Linguistics  
1611 North Kent Street  
Arlington, Virginia 22209

I will address my comments to the section dealing with language development, since this is the area with which I feel most comfortable.

I am very much in agreement with the statements made in the language development section of the interim report; however, I feel that, regarding the Spanish-speaking child, only one side of the coin was considered. To the Spanish-speaking population being served by Head Start, the phrase "to be able to deal with his environment and later responsibilities in school and life" means something different from what it means for monolingual English-speaking children. There is the issue of children with a language, a way of expressing themselves, a culture that is different from the dominant Anglo culture. To ignore this in 15% of Head Start sites is to put aside a major responsibility of Head Start.

The material given us seems to indicate that Head Start is interested only in developing certain English proficiencies in the participants. This appears to contradict the concept of preparing the child to be able to deal with his environment, for if the child has Spanish as a first language and lives in a neighborhood (barrio) where Spanish is the primary vehicle for communication, Head Start is being derelict in its duty by

discouraging the development of his first language. Moreover, a Head Start program that wastes valuable time in teaching a Spanish-speaking child English instead of beginning instruction and introducing concepts in Spanish, does not appear to have the Spanish-speaking children's best interests at heart. Aside from this, there is the child's self-concept to consider. The present evaluation design, by not even considering Spanish proficiency, is declaring it worthless.

The issue that Head Start must resolve is similar (if not identical) to that confronting Title VII bilingual programs, namely, should it attempt to move the children away from their first language, should they teach English as a Second Language without any regard for the child's psychological being and with no regard for maintaining the first language, or should such programs be what they profess to be -- truly bilingual -- with equal emphasis on the development of both the native and the second language? If, as is stated in the document on page 45, "language is in many ways a central issue in the development of social competence," it would seem that the same principle would apply in the case of children having more than one language, i.e., that both languages would be equally "central" in the development of a child's social competence.

As for the assessment of the effect of Head Start on the language of Spanish-speaking children, there is a dire need for

basic research on the language of Spanish-speaking populations. Without such exploratory studies, it will be difficult if not impossible to determine what is average for a certain age, Spanish-speaking group (Cuban, Puerto Rican, Chicano), or sociolinguistic context. Language acquisition studies in the United States to date have focused almost exclusively on English; only one study has focused on the acquisition of syntactic structures in Spanish to my knowledge (my own). I would suggest that if Head Start is really serious about serving the needs of the Spanish-speaking population, it should invest in research which would provide information on the language acquisition process of Spanish-speaking children; the influence of English on their Spanish as they get older; the influence of peers, siblings, and parents; the social contexts in which children fluent in both languages use each language, and in which contexts they code-switch (mix the languages); which teaching strategies are most effective in producing the desired outcomes; the most appropriate teacher training for teachers with significant numbers of Spanish speakers in their classrooms.

From the above basic research would come some indication of what is average for the different Spanish-speaking populations. This would allow for the construction of assessment instruments, normed and validated on the population that it was meant to serve and not an Anglo norm. For all the different dimensions

and aspects of language mentioned in the language development section, it is necessary to have a Spanish counterpart (in addition to the often-mentioned English goals). In order to prepare adequate measures of the Spanish-speaking child's attainment in the two languages, research is needed. I cannot emphasize that enough.

To cover each of the specifics raised in the evaluation design regarding language development, a more extensive paper would be required. Vera John Steiner's paper can certainly not be responded to adequately in a format such as this. Hopefully an opportunity to further amplify the above remarks will be provided to me.

- ① Oral Language Dominance Measure (Spanish, English)  
El Paso Public Schools (Vocabulary, Comprehension, Controlled Conversation)  
El Paso, Texas
- ② The Peter James Test of Spanish-English Language Dominance  
Mr. Peter James  
ILSC, University of Texas, Austin
3. Oral Language Ability Tests, K-3, English  
(Performance Objectives, Pilot Project, Oral Language)  
Texas Child Migrant Program (Vocabulary, Comprehension, Rhetoric)  
Texas Education Agency, Austin
4. Test for Auditory Comprehension of Language (English/Spanish),  
developed by Elizabeth Carrow, Ph.D. (Vocabulary, Morphology, Grammar, Syntax)  
Urban Research Group  
301 West 16th  
Austin, Texas 78701
- ⑤ HABLA (Spanish-English)  
Abernathy Title VII Project  
Abernathy, Texas 79311 (Vocabulary)
- ⑥ Dos Amigos Verbal Language Scales (Spanish-English)  
Professor Donald E. Critchlow (Vocabulary)  
Texas A & I University  
Laredo, Texas
- ⑦ Bilingual Test of Language Dominance (Spanish-English),  
prepared by the Arizona State Department of Education  
Southwest Research Associates, Inc. (Word meanings, word associations)  
Albuquerque, New Mexico
8. Oral Language Assessment "OLA", Diagnostic (English only)  
John R. Munden (Syntax)  
Reading Center  
Arizona State University  
Tempe, Arizona
9. Oral Bilingual Test, Gloria and David - English-Spanish  
Mr. Gib Devine (Pronunciation, Phonology, Listening Comprehension)  
Language Arts, Inc.  
Austin, Texas
10. Tests of Grammatically Correct Spanish and English, Las  
Cruces Bilingual Education Project (Grammar)  
Bilingual Project Coordinator  
301 West Amador Avenue  
Las Cruces, New Mexico 88001
11. Inter-American Series, by Herschel T. Manuel (English-Spanish)  
Guidance Testing Associates (Comprehension)  
6516 Shirley Avenue  
Austin, Texas 78752
12. Poabody Picture Vocabulary Test (Spanish, Portuguese, English)  
American Guidance Service (Word Meanings)  
Publishers Building  
Circle Pines, Minnesota 55014
13. Screening Test for State Bilingual Education Program, developed  
by Division of Instruction (English only)  
Corpus Christi Independent School Dist.  
Corpus Christi, Texas (Comprehension, Vocabulary, Verbal Expression, Pronunciation)

# Texas Education Agency -17-



• STATE BOARD OF EDUCATION

• STATE COMMISSIONER OF EDUCATION

• STATE DEPARTMENT OF EDUCATION

201 East 11th Street

Austin, Texas

78701

February 12, 1974

Dr. Ramón García  
Research Associate  
Department of Health, Education,  
and Welfare  
Office of the Secretary  
P. O. Box 1182  
Washington, D.C. 20013

Dear Ramón:

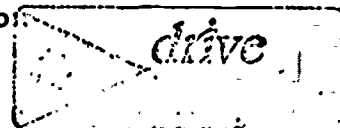
Enclosed you will find a reaction paper to the proposal and interim report on the evaluation of social competencies in Head Start children. I sincerely hope that the information will be useful to you and to the Rand Corporation personnel in the refinement of the interim report.

In the interest of clarity, I hope that the internal policy matters such as the use to be made of the results and their implications for policy as it relates to the Spanish-speaking children will be resolved. Also I would hope that Rand Corporation will provide a group such as the one of which I was a member to review proposals and/or matters which affect the Spanish-speaking population.

Ramón, I apologize for the lateness of my report. If I may be of service to you, please do not hesitate to call on me.

Sinceramente,

Arturo Luis Gutiérrez, Ph.D.  
Director, Special Programs  
Office of International and Bilingual Education





REACTION PAPER  
Arturo Luis Gutierrez, Ph.D.

The proposal for Designing an Evaluation of Social Competency in Head Start Children is certainly an ambitious one. Clearly, the list of specific elements to be considered in the development of the technical design on page 7 of the proposal speaks to the identification of independent variables that could influence evaluation results. It is indicated that "cultural background of the participants" is an example of the type of variable that will be considered in the development of the design.

In accordance with your request, following are my observations after a review of the proposal and the interim report:

Implication of these papers for the Spanish-speaking child

- Neither the culture of the Spanish-speaking nor its most central manifestation, language, with all its varieties, appears to have been taken into consideration in the development of the design.
- A base battery to measure social competencies administered to all participating children would be inadequate because it assumes at least a certain degree of English language competence for all children and does not provide for cultural differences.
- The results of a base battery that measures social competencies administered to all participants without regard to language and culture of the children could

have negative effects on the total Head Start program.

- . The results of a base battery that measures social competencies administered to all participants without regard to local program philosophy, program interpretation, staffing pattern, baseline data of children, teacher focus, teacher-pupil classroom interaction, etc. would provide uninterpretable results.
- . It is recommended that the following person be contacted regarding the Townsend's Bilingual Interaction Analysis System (BIAS) to determine the interaction patterns of bilingual teachers who teach in Spanish and English.

Dr. Darryl T. Townsend  
Education Annex F38  
The University of Texas at Austin  
Education Administration Department  
Austin, Texas 78712

Conditions that must be met before an evaluation of social competencies can be conducted on Spanish-speaking children

- . Determine language dominance of children at beginning of program.
- . Determine the degree (length of time) to which one language or the other or both is used for instruction during the day (in the program).
- . Determine teacher emphasis either on first (Spanish) or second (English) language development, social-emotional development or other areas.

- . Determine whether a structured language program (Spanish or English as a Second Language) has been implemented.
- . Determine cultural and language make-up of the participants in the Head Start room.
- . Determine the overall classroom environment for conduciveness to the development of social competencies.

Social competencies that Spanish-speaking children can be expected to acquire as a result of Head Start

There do not appear to be any differences in the competencies that Spanish-speaking children can be expected to acquire as a result of Head Start in comparison with other Head Start participants, except, perhaps, in the area of language itself and its positive or negative effects on the child. If a child is dominant in Spanish, the use of that language for instruction, the teacher's attitude toward the use of the language, the teacher's acceptance or rejection of the language, values, total cultural make-up of the child, could have serious negative effects on the child's personality -- his self-concept, confidence, his perceptions, etc.

What the statement "to be able to deal with his environment and later responsibilities in school and life" means for Spanish-speaking children

- . I feel that Head Start probably has or should have its greatest impact in preparing children for success in school without alienating him from his family and his home environment. Head Start should play a major role in the development of the child as a truly bilingual-multicultural individual.

Be prepared to discuss specific sections of the Report (health, language, socio-emotional), as per your interest and expertise

- . Nutrition

- .. It is recommended that the evaluators check with State Health Departments to find out if there is any research on stress-related diseases.
- .. Look at the 1968 research on Head Start related to stress-related diseases (accidents) to see if it is cross-referenced by sub-population groups.

- . Motor-Perceptual Development

- .. Page 33-#1(f) Auditory - It is not clear what discrimination and articulation mean in this context. However, if we are talking about the child's articulation of language sounds, the

child's language, whether it is Spanish or English, must be taken into consideration.

- .. Page 33 #2 Classification - The following are suggested as possibilities for the development of measures of classification.

Dr. Celia Stendler Lovatelli  
c/o American Science and Engineering  
Education Division  
20 Overland Street  
Boston, Massachusetts

or

Professor of Elementary Education  
Department of Education  
University of Illinois  
Chicago, Illinois 60680

Dr. Edward de Avila  
Bilingual Children's Television  
The Ordway Building, Suite 2350  
2150 Valdez Street  
Oakland, California 94612

Both of the above named have done extensive work on Piagetian tasks and their evaluation.

#### Language

- .. I am enclosing a list of language measures that are being used throughout the Southwest with Spanish-speaking children. Some of these may have merit for use as they are or with further refinements.
- .. At the risk of beating a dead horse to death, language testing must be adjusted to the child's dominant language and to the type of program implemented e.g. bilingual versus all English versus all Spanish; Spanish oral language development versus English as a second language; structured versus non-structured;

teacher versus teacher aide versus parent; etc.

- .. Page 43 - The Language Children Use - Probably the closest assessment instrument that we have and that can be adapted to Spanish is the Oral Language Ability Test developed by the Texas Child Migrant Program, Texas Education Agency, 201 East 11th Street, Austin, Texas.

Appendix E

TECHNICAL INFORMATION ON MEASURES CONSIDERED  
FOR PERCEPTUAL-MOTOR/COGNITIVE/LANGUAGE BATTERY

This appendix includes:

CIRCUS battery

Description of national samples

Reliability statistics

Standard errors of measurements

Description of individual tests

Other measures considered

Table: Correlations on total and subscale scores and background  
variables



Description of CIRCUS National Samples  
(Based on largest N - What Words Mean)

	<u>Kindergarten</u> Total N=1930		<u>Nursery School</u> Total N=946
<u>Age</u>			
<5.0	180 (10%)	<4.3	91 (9%)
5.0 - 5.3	367 (20%)	4.3 - 4.6	168 (18%)
5.3 - 5.6	449 (24%)	4.6 - 4.9	211 (23%)
5.6 - 5.9	432 (23%)	4.9 - 5.0	208 (23%)
5.9 - 6.0	316 (17%)	5.0 - 5.3	166 (18%)
>6.0	99 (5%)	>5.3	72 (8%)
<u>Region</u>			
North Atlantic	552 (29%)		276 (29%)
Great Lakes	532 (28%)		258 (27%)
Southeast	431 (22%)		215 (23%)
W/Southwest	415 (22%)		197 (21%)
<u>Sex</u>			
Boys	990 (51%)		484 (51%)
Girls	940 (49%)		462 (49%)
<u>Race</u>			
White	1587 (82%)		819 (87%)
Black	227 (12%)		108 (11%)
Spanish	68 (4%)		12 (1%)
Other	48 (2%)		7 (1%)
<u>Occupation</u>			
Professional	477 (25%)		420 (44%)
Skilled	975 (51%)		381 (40%)
Unskilled or	478 (25%)		145 (15%)
Unemployed			
<u>Previous School</u>			
Yes	650 (34%)		385 (41%)
No	1280 (66%)		561 (59%)

# STATISTICS ON ADMINISTRATION OF CIRCUS

## Kindergarten

Test	Number of Items	Kindergarten				Nursery School				Split Half
		N	Mean	S.D.	Alpha	N	Mean	S.D.	Alpha	
I. What Words Mean	40	1930	30.1	5.9	.83	946	27.8	6.7	.86	.87
II. How Much and How Many	40	1883	30.5	6.2	.86	934	28.0	7.1	.87	.87
III. Look-a-likes	26	563	20.7	4.4	.84	250	19.1	4.9	.84	.88
IV. Copy What You See (maximum score 42)	14	572	33.9	5.8	.87	267	31.1	8.2	.90	.90
V. Finding Letters and Numbers	20	546	14.2	4.6	.86	290	15.5	4.2	.86	.86
VI. Noises	24	563	19.2	3.1	.71	286	19.0	3.9	.81	.80
VII. How Words Sound	44	644	39.4	5.1	.87	300	35.8	6.7	.91	.91
VIII. How Words Work	26	594	20.8	3.3	.71	252	20.4	3.8	.78	.81
IX. Listen to the Story	24	621	18.9	4.2	.79	269	18.0	4.2	.77	.78
X. Say and Tell Part II.	44	590	59.0	15.7	.91	227	57.0	15.1	.90	.93
XI. Do You Know	32	591	28.2	3.5	.77	286	26.5	4.2	.79	.78
XII. See and Remember	20	568	15.3	3.2	.71	260	14.7	3.1	.68	.61
XIII. Think it Through	32	600	22.2	5.4	.81	273	21.5	5.7	.82	.82

STANDARD ERRORS OF MEASUREMENT

<u>Instrument</u>	<u>Kindergarten</u>	<u>Nursery School</u>
I. What Words Mean	2.39	2.55
II. How Much and How Many	2.36	2.54
III. Look-a-likes	1.80	1.94
IV. Copy What You See	2.10	2.56
V. Finding Letters and Numbers	1.70	1.57
VI. Noises	1.70	1.70
VII. How Words Sound	1.81	2.03
VIII. How Words Work	1.74	1.80
IX. Listen to the Story	1.93	2.00
X. Say and Tell (Part II)	4.80	4.82
XI. Do You Know	1.68	1.93
XII. See and Remember	1.73	1.78
XIII. Think It Through	2.35	2.39

Test: CIRCUS NO. 1 - "WHAT WORDS MEAN"\*

Outcome: Semantics

Purpose: To assess receptive vocabulary

Task: Associating nouns, verbs, and modifiers with pictures

Age range: Preschool and kindergarten

Administration: Group; basic sample

Costs: \$3.75 per package of 10 booklets; \$.70 per measure scored

Norms: Age, region, sex, race, parent's occupation, previous school

Reliability: Alpha = .86; split-half r = .87

Validity: Not available

Comments: o Must be screened for unnecessary penalties to various ethnic groups

---

\*Test includes 40 items: nouns, 20 items; verbs, 12 items; modifiers, 8 items.

**Test:** CIRCUS NO. 2 - "HOW MUCH! AND HOW MANY"\*

**Outcome:** Numerical readiness

**Purpose:** To measure competence in quantitative concepts

**Task:** Demonstrating understanding of enumeration, counting, one-to-one correspondence, ordination, comparison, quantitative language, etc., through identification of appropriate pictures

**Age range:** Preschool and kindergarten

**Administration:** Group; basic sample

**Costs:** \$3.75 per package of 10 booklets; \$.70 per measure scored

**Norms:** Age, region, sex, race, parent's occupation, previous school

**Reliability:** Alpha = .87; split-half r = .87

**Validity:** Not available

**Comments:**

---

\*Test includes 40 items: counting, 12 items; relational terms, 14 items; numerical concepts, 14 items.

Test: CIRCUS NO. 3 - "LOOK-ALIKES"\*

Outcome: Visual recognition and discrimination

Purpose: To measure visual discrimination ability

Task: Matching identical letters, numbers, drawings

Age range: Preschool and kindergarten

Administration: Group; basic sample

Costs: \$3.75 per package of 10 booklets; \$.70 per measure scored.

Norms: Age, region, sex, race, parent's occupation, previous school

Reliability: Alpha = .84; split-half  $r = .88$

Validity: Not available

Comments:

---

\*Test includes 26 items: reversals, 11 items; complex matching, 13 items.

**Test:** CIRCUS NO. 4 - "COPY WHAT YOU SEE"\*

**Outcome:** Fine motor skills (visually guided)

**Purpose:** To measure perceptual-motor coordination

**Task:** Copying letters and numbers

**Age range:** Preschool and kindergarten

**Administration:** Group; basic sample

**Costs:** \$3.75 per package of 10 booklets

**Norms:** Age, region, sex, race, parent's occupation,  
previous school

**Reliability:** Alpha = .90; split-half r = .90

**Validity:** Not available

**Comments:**

- o Also taps persistence and attention (indices of school readiness)
- o No score for reversals

---

\*Test includes 15 items.



Test: CIRCUS NO. 5 - "FINDING LETTERS AND NUMBERS"\*

Outcome: Preliteracy skills, numerical readiness

Purpose: To assess letter-number recognition

Task: Identifying capital letters, lower-case letters,  
and numbers

Age range: Preschool and kindergarten

Administration: Group

Costs: \$3.75 per package of 10 booklets; \$.70 per measure  
scored

Norms: Age, region, sex, race, parent's occupation, previous  
school

Reliability: Alpha = .86; split-half r = .86

Validity: Not available

Comments:

---

\*Test includes 20 items: capital letters, 9 items; lower-case  
letters, 6 items; numbers, 5 items.

Test: CIRCUS NO. 6 - "NOISES"\*

Outcome: Auditory recognition skills

Purpose: To measure ability to discriminate real-world sounds

Task: Associating taped sounds with pictures

Age range: Recommended for prekindergarten only

Administration: Group

Costs: \$4.25 per package of 10 booklets; \$.70 per measure scored

Norms: Age, region, sex, race, parent's occupation, previous school

Reliability: Alpha = .81; split-half r = .80

Validity: Not available

Comments: 

- o Lots of fun for children
- o Gains unlikely

---

\*Test includes 24 items.

(DISCARDED)

**Test:** CIRCUS NO. 7 - "HOW WORDS SOUND"\*

**Outcome:** Linguistic competence

**Purpose:** To assess auditory discrimination ability (phonemes)

**Task:** Associating words having similar phonemes with pictures representing the words

**Age range:** Preschool and kindergarten

**Administration:** Group

**Costs:** \$3.75 per package of 10 booklets; \$.70 per measure scored

**Norms:** Age, region, sex, race, parent's occupation, previous school

**Reliability:** Alpha = .91; split-half r = .91

**Validity:** Not available

**Comments:**

---

\*Test includes 44 items.

(DISCARDED)

**Test:** CIRCUS NO. 8 - "HOW WORDS WORK"\*

**Outcome:** Linguistic competence

**Purpose:** To assess child's receptive functional language

**Task:** Discriminating (through pictures) between verb forms, prepositions, negatives and positives, and sentence orders

**Age range:** Preschool and kindergarten

**Administration:** Group

**Costs:** \$3.75 per package of 10 booklets; \$.70 per measure scored

**Norms:** Age, region, sex, race, parent's occupation, previous school

**Reliability:** Alpha = .78; split-half  $r = .81$

**Validity:** Not available

**Comments:**

---

\*Test includes 26 items: verb forms, 8 items; prepositions/negation/conjunctions, 10 items; syntax, 8 items. Total score used for group purposes only; not reported for individual children.

(DISCARDED)

**Test:** CIRCUS NO. 9 - "LISTEN TO THE STORY" \*

**Outcome:** Semantics

**Purpose:** Test of comprehension, interpretation and recall of oral language

**Task:** Understanding a story

**Age range:** Preschool and kindergarten

**Administration:** Group; basic sample

**Costs:** \$3.75 per package of 10 booklets; \$.70 per measure scored

**Norms:** Age, region, sex, race, parent's occupation, previous school

**Reliability:** Alpha = .77; split-half  $r = .78$

**Validity:** Not available

**Comments:**

- o Should be encouraged in bilingual settings
- o Spanish version should also be developed, tapping semantic competence in native as well as second language

---

\*Test includes 25 items: comprehension, 15 items; interpretation, 10 items.

Test: CIRCUS NO. 10 - "SAY AND TELL" \*

Outcome: Competence in language use

Purpose: To assess ability, under changing conditions, to produce oral language

Task: Three parts: description of common objects, grammatical use of language, telling a story

Age range: Preschool and kindergarten

Administration: Individually administered; basic and subsamples

Costs: \$3.75 per package of 10 booklets

Norms: Age, region, sex, race, parent's occupation, previous school

Reliability: Alpha = .90; split-half  $r = .93$

Validity: Not available

Comments:

---

\*Test includes 66 items: description, 16 items; productive functional language, 38 items (plurals, 9; verbs, 9; subject-verb agreement, 12; comparisons, 8); narration (quality), 12 items.

**Test:** CIRCUS NO. 11 - "DO YOU KNOW"\*

**Outcome:** Cognitive aspects of social competence

**Purpose:** To assess general information

**Task:** Demonstrating knowledge about health and safety, physical and social environments, consumer concepts, recreation, etc.

**Age range:** Preschool and kindergarten

**Administration:** Group

**Costs:** \$3.75 per package of 10 booklets; \$.70 per measure scored.

**Norms:** Age, region, sex, race, parent's occupation, previous school

**Reliability:** Alpha = .79; split-half  $r = .78$

**Validity:** Not available

**Comments:**

---

\*Test includes 32 items.

(DISCARDED)

00101



**Test:** CIRCUS NO. 12 - "SEE AND REMEMBER"\*

**Outcome:** Active memory, cognitive processes

**Purpose:** To assess visual and associative memory

**Task:** Remembering pictures and name associations,  
both immediately and after intervening experiences

**Age range:** Preschool and kindergarten

**Administration:** Group

**Costs:** \$3.75 per package of 10 booklets; \$.70 per measure  
scored

**Norms:** Age, region, sex, race, parent's occupation,  
previous school

**Reliability:** Alpha = .68; split-half r = .61

**Validity:** Not available

**Comments:**

---

\*Test includes 20 items.

(DISCARDED)

Test: CIRCUS NO. 13 - "THINK IT THROUGH" \*

Outcome: Classification skills, problem-solving

Purpose: To assess problem-solving ability

Task: Identifying problems, classifying drawings, identifying first event in a sequence, evaluating problem solution

Age range: Preschool and kindergarten

Administration: Group; basic sample

Costs: \$3.75 per package of 10 booklets; \$.70 per measure scored

Norms: Age, region, sex, race, parent's occupation, previous school

Reliability: Alpha = .82; split-half  $r = .82$

Validity: Not available

Comments: o Also taps sequencing ability

---

\*Test includes 32 items: problem identification, 6 items; classification, 17 items; solution evaluation and time sequence, 9 items.

**Test:** CIRCUS NO. 14 - "MAKE A TREE"

**Outcome:** Originality, flexibility

**Purpose:** To assess divergent pictorial production

**Task:** Constructing a tree with gummed labels (performed twice)

**Age range:** Preschool and kindergarten

**Administration:** Group

**Costs:** \$3.75 per package of 10 booklets; scored locally

**Norms:** Age, region, sex, race, parent's occupation, previous school

**Reliability** Not available

**Validity:** Not available

**Comments:**

- o Not a measure of creativity
- o Correlates well with verbal measures
- o Does not require fine motor skills

(DISCARDED)

**Test:** CIRCUS BEHAVIOR INVENTORY - NO. 16

**Outcome:** On-task/off-task behavior; anxiety in achievement situation

**Purpose:** To assess child's reaction to testing situation

**Task:** Interest, attention, and other aspects of child's reaction to CIRCUS measures

**Age range:** Preschool and kindergarten

**Administration:** Rating made by teacher (or independent observer); group; basic and subsamples

**Costs:** \$1.00 per package of 10 booklets

**Norms:** Age, region, sex, race, parent's occupation, previous school

**Reliability:** Not available

**Validity:** Not available

**Comments:**

- o Indirect child measure
- o Suggested that this behavior should also be observed in other testing situations

**Test:** BENDER-GESTALT TEST (Original and five modifications)

**Outcome:** Perceptual/motor functioning

**Purpose:** Clinical diagnostic instrument (e.g., brain damage)

**Task:** Copying designs

**Age range:** 4 years and older

**Administration:** Individual; 10 minutes

**Costs:**

**Norms:**

**Reliability:** Low retest reliability

**Validity:** Low predictive validity

**Comments:**

(DISCARDED)

Test: BOEHM TEST OF CONCEPTS

Outcomes: School readiness

Purpose: To measure mastery of concepts considered necessary for elementary school achievement

Task: 50 concept-picture matching items

Age range: K, 1st, 2nd grades

Administration: 30-40 minutes in 2 sessions; individually or in group; simple to administer (designed for teachers)

Costs:

Norms: By grades and SES; standardization sample not further specified

Reliability: Split-half  $r = .12-.94$ ; alternate-form  $r = .55-.92$

Validity: Nothing more than face validity provided

Comments:

- o Alternate form available for retesting
- o Diagnostic, screening device
- o UCLA critique: confounding of concepts with verbal labels, i.e., comprehension of English; essentially a vocabulary test; often ambiguous
- o Too easy for 2nd graders

(DISCARDED)

**Test:** COATES' PRESCHOOL EMBEDDED FIGURES TEST

**Outcome:** Information-gathering strategies

**Purpose:** Measures field independence, analytic functioning

**Task:** 27 drawings; in each is embedded a triangle

**Age range:** 3 years and older

**Administration:** Difficult to administer at younger ages;  
10-15 min.; individual

**Costs:**

**Norms:**

**Reliability:** No conclusive evidence

**Validity:** Low concurrent validity (.00-.36)

**Comments:**

- o Sensitive to age, sex and cultural background
- o Personality correlates found

(DISCARDED)



**Test:** ETS ENUMERATION TEST II

**Outcomes:** Numerical readiness

**Purpose:** To measure place required to learn concept of number

**Task:** 21 items, 4 subtests: counting, touching, same number matching, same order matching

**Age range:** 2-1/2 to 4 years old

**Administration:** Individual; simple

**Costs:**

**Norms:** Age, sex, preschool experience, ethnicity

**Reliability:** Test-retest  $r = .028-.946$ ; internal consistency  $= .77$ ;  
KR-20  $= .72-.79$

**Validity:** Concurrent  $= .01-.69$  (PSI)

**Comments:**

- o Scores sensitive to age, sex, SES
- o Its intercorrelations very low
- o Low correlation among subtests
- o Loads highly on "g"; no factor analytic evidence for underlying specific numerical skill

(DISCARDED)

Test: ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES

Outcome: Ability to express oneself verbally

Purpose: Diagnostic test of cognitive functioning: language, perception, and short-term memory

Tasks: 12 subtests: visual and auditory input -- vocal and motor output

Age range: 2-10 years

Administration: Individual; 45-60 minutes

Costs: \$43.50/set of test materials

Norms: Standardized on 962 "normal" children of average IQ from middle-class communities in Midwest

Reliability: Moderate; retest  $r = .12 - .86$

Validity:

Comments:

- o Complicated to administer
- o Usefulness of norms questionable
- o Unsuitable for lower class and minority children unless their progress toward middle-class language norms sought

(DISCARDED)

Test: MCCARTHY SCALE OF CHILDREN'S ABILITIES

Outcome: Motor skills, information processing, school readiness

Purpose: Individual test of general mental ability

Taks: 18 tests, six scales: Verbal, perceptual-performance, quantitative, general cognitive, memory, motor

Age range: 2-1/2 through 8-1/2 years

Administration: Special training and instruction required; scoring complex, time-consuming, and subjective; 45-60 min.;

Costs: Expensive

Norms: By sex, race (16% nonwhite), geographic region, father's occupation

Reliability: Split-half r for 6 scales = .79-.93; test-retest r = .69-.91

Validity: Good concurrent validity (SB, WPPSI, MAT); low predictive validity

Comments:

- o Parallels the Stanford-Binet - both measuring general rather than specific competencies
- o Specific item of perceptual-performance subtest may be chosen: copying a sequence tapped by examiner on a xylophone to assess auditory-visual integration skills
- o Relatively few black-white differences on scales encourages use with disadvantaged

(DISCARDED)

**Test:** METROPOLITAN READINESS TESTS

**Outcome:** Reading readiness

**Purpose:** To measure skills necessary for readiness in 1st grade instruction

**Tasks:** 102 items: picture vocabulary, comprehension, matching, alphabet, numbers, copying

**Age range:** K-1st grade

**Administration:** 60 min., maximum group of 15 pupils

**Costs:**

**Norms:** "Representative" national sample of school children

**Reliability:** Odd-even  $r=.91-.94$ ; alternate form  $r=.91$

**Validity:** Questionable content validity, moderate congruent validity, limited predictive validity

**Comments:**

- o Lack of useful, diagnostic information
- o To large extent, an achievement test
- o Relevancy to Spanish-speaking populations highly questionable

(DISCARDED)

**Test:** PEABODY PICTURE VOCABULARY TEST

**Outcomes:** Semantics, school readiness

**Purpose:** To measure verbal intelligence by measuring receptive vocabulary

**Task:** 150 stimulus word-picture matching items

**Age range:** 2-1/2 to 18 years old

**Administration:** Individual; not timed, but approx. 15 min.; no special training required, but scoring difficult

**Costs:**

**Norms:** Originally white children in and around Nashville; now available on HSPV sample; lacking for Spanish-speaking populations

**Reliability:** Alternative form  $r = .67-.84$

**Validity:** Content, low predictive, concurrent  $r = .04-.91$

**Comments:**

- o High loading on "g"
- o Confounding of score with attention span
- o Six months of age can create big jumps or drops in score
- o Not substitute for IQ test, although IQ score is calculated

(DISCARDED)

Test: PRESCHOOL INVENTORY

Outcome: School readiness

Purpose: To measure achievement in areas regarded as necessary for success in school

Task: 32, 64, or 85 items on general knowledge, listing and word meanings, listening and comprehension, writing and copying, quantitative concepts, speaking and labeling

Age range: Three to six years

Administration: Approximately 20 minutes; individual

Costs:

Norms: Head Start children, age, region, preschool experience, race, sex

Reliability: High; KR-20 = .91; split-half  $r = .92$

Validity: Good concurrent validity; face validity

Comments:

- o Designed for diagnostic purposed rather than test of generalized cognitive ability
- o Not "culture-fair"; reflects biases of school
- o May have strong practice effects

(DISCARDED)

**Test:**     SEQUIN FORM BOARD

**Outcomes:**       Visual recognition and discrimination

**Purpose:**         To assess visual discrimination and matching, and eye-hand coordination

**Task:**           Timed placing of shaped blocks in corresponding recesses on board

**Age range:**

**Administration:** Simple; individual; requires tester practice

**Costs:**

**Norms:**         Middle and lower SES

**Reliability:**    Test-retest  $r = .60-.74$

**Validity:**       Unclear from wide range of correlations with other timed or matching tasks

**Comments:**

- o Significant age, sex, SES (although small) differences
- o Loading on "g"
- o Questionable whether measure of visual-motor integration and motor speed rather than visual discrimination
- o Taps different outcomes at different ages: ability to comprehend instructions in younger, and abilities of spatial judgments and relations in older child

(DISCARDED)



**Test:** SIGEL OBJECT CATEGORIZATION TEST

**Outcomes:** Perceptual processes, recognition, classification

**Purpose:** To measure classification abilities of young children

**Task:** Active and passive sorting of 12 objects/pictures

**Age range:**

**Administration:** Requires extensive training; scoring difficult

**Costs:**

**Norms:**

**Reliability:** Alpha =.44-.91; test-retest r=.06-.26

**Validity:** Predictive, concurrent

**Comments:**

- o Fairly difficult task
- o Significant SES differences
- o Difficult to use in large-scale evaluations
- o "Provides a wealth of data" (ETS) but nature of data unspecified
- o Allows child to manipulate objects rather than paper and pencil task

(MIGHT BE CONSIDERED FOR SPECIAL STUDY)

Test: STANFORD-BINET INTELLIGENCE SCALE (L-M form)

Outcome: General intelligence; mental adaptability

Purpose: To assess general intelligence

Task: Age scale

Age range: 2 years and older

Administration: Individual; 30-90 minutes

Costs: \$45/set of test materials; \$4.65/35 booklets

Norms: Standardized on white sample

Reliability: Correlation coefficients: .83-.98

Validity: High predictive validity

Comments:

- o Attempts to measure underlying unitary concept of intelligence - singular index
- o Serious question of cultural and socio-economic bias
- o Heavy verbal loading

(DISCARDED)

Test: WECHSLER PRESCHOOL AND PRIMARY SCALE OF INTELLIGENCE

Outcome: Perceptual/motor processes

Purpose: To assess general intelligence

Tasks: 11 verbal and performance subtests

Age range: 4 to 6 1/2 years

Administration: Individual; 50-75 minutes

Costs: \$26/set of testing materials

Norms: Standardized on 100 boys and 100 girls in each of six half-year age groups; stratified by geographic region, urban-rural, father's occupation, color

Reliability: Test-retest  $r = .86 - .92$ ; split-half  $r = .93 - .96$

Validity: Concurrent:  $.58$  (PPVT) -  $.75$  (Stanford-Binet); good predictive validity

Comments:

- o Comparison with age peers only
- o Compact and attractive test materials
- o Subjectivity in scoring of some subtests
- o Unduly long for preschool children

(DISCARDED)

Table E1

CORRELATIONS ON TOTAL AND SUBSCALE SCORES AND BACKGROUND VARIABLES  
Full Correlation Matrix -- CIRCUS Nursery School Data

Measure	Receptive Vocabulary	Quantitative Concepts	Visual Discrim.	Perceptual Motor Coordination	Letter and Number Discrim.	Listening Comprehension	Productive Functional Language	Parts IA & IB	Quality	Part II	Problem-solving
NURSERY SCHOOL											
Receptive Vocabulary (CIRCUS No. 1)	1.00	0.58	0.45	0.34	0.36	0.46	0.32	0.36	0.03	0.31	0.43
Quantitative Concepts (CIRCUS No. 2)	0.58	1.00	0.69	0.53	0.55	0.63	0.52	0.44	0.07	0.51	0.66
Visual Discrimination (CIRCUS No. 3)	0.45	0.69	1.00	0.43	0.50	0.65	0.14	0.32	0.10	0.09	0.69
Perceptual-Motor Coordination (CIRCUS No. 4)	0.34	0.53	0.43	1.00	0.43	0.60	0.06	-0.04	0.01	0.08	0.47
Letter and Number Discrimination (CIRCUS No. 5)	0.36	0.55	0.50	0.43	1.00	0.18	0.40	0.14	0.13	0.41	0.31
Listening Comprehension (CIRCUS No. 9)	0.46	0.63	0.65	0.60	0.18	1.00	0.53	0.51	0.19	0.53	0.67
Productive Functional Language (CIRCUS No. 10)	0.32	0.52	0.14	0.06	0.40	0.53	1.00	0.66	0.37	0.97	0.32
Parts IA and IB	0.36	0.44	0.32	-0.04	0.14	0.51	0.66	1.00	0.31	0.52	0.38
Quality	0.03	0.07	0.10	0.01	0.13	0.19	0.37	0.31	1.00	0.18	0.38
Part II	0.31	0.51	0.09	0.08	0.41	0.53	0.97	0.52	0.18	1.00	0.13
Problemsolving (CIRCUS No. 13)	0.43	0.66	0.69	0.47	0.31	0.67	0.32	0.38	0.38	0.13	1.00
Age	0.23	0.27	0.22	0.24	0.11	0.23	0.25	0.21	0.04	0.24	0.26
Sex	0.00	0.00	-0.02	0.03	0.01	0.10	-0.02	-0.09	0.03	-0.01	0.02
White	0.09	0.20	0.20	0.11	0.22	0.23	0.27	0.05	0.07	0.29	-0.04
Black	0.08	-0.16	-0.19	-0.12	-0.22	-0.13	-0.25	-0.01	-0.10	-0.26	-0.00
Spanish	0.00	-0.08	0.0	-0.02	-0.06	-0.13	0.02	0.04	0.12	-0.00	0.08
Other	-0.08	-0.10	-0.10	0.05	0.00	-0.21	-0.12	-0.12	-0.02	-0.12	0.01
North Atlantic	-0.02	-0.04	-0.11	-0.02	-0.07	0.00	0.08	0.06	0.08	0.07	-0.23
Great Lakes	0.10	0.07	0.05	0.08	0.03	-0.03	-0.01	0.10	-0.14	-0.00	0.18
Southeast	-0.08	-0.07	-0.02	-0.05	-0.09	0.10	-0.02	0.03	-0.08	-0.01	-0.05
West/Southwest	-0.00	0.04	0.08	-0.01	0.14	-0.08	-0.05	-0.19	0.14	-0.06	0.12
Professional	0.05	0.21	0.28	-0.05	0.30	0.21	0.03	-0.02	0.11	0.02	0.24
Skilled	-0.00	-0.05	-0.05	-0.09	-0.09	0.18	0.07	0.01	-0.03	0.08	-0.03
Unskilled	-0.07	-0.22	-0.30	-0.05	-0.34	-0.03	-0.15	0.00	-0.12	-0.15	-0.26
Previous School	0.06	0.15	0.20	0.10	-0.16	0.21	0.15	0.19	0.23	0.09	0.30

Table E1

CORRELATIONS ON TOTAL AND SUBSCALE SCORES AND BACKGROUND VARIABLES  
Full Correlation Matrix -- CIRCUS Kindergarten Data

Measure	Receptive Vocabulary	Quantitative Concepts	Visual Discrim.	Perceptual Motor Coordination	Letter and Number Discrim.	Listening Comprehension	Productive Functional Language	Parts IA & IB	Quality	Part II	Problem-solving
KINDERGARTEN											
Receptive Vocabulary (CIRCUS No. 1)	1.00	0.68	0.46	0.44	0.49	0.70	0.51	0.38	0.12	0.50	0.65
Quantitative Concepts (CIRCUS No. 2)	0.68	1.00	0.60	0.53	0.62	0.70	0.44	0.35	0.12	0.43	0.68
Visual Discrimination (CIRCUS No. 3)	0.46	0.60	1.00	0.50	0.47	0.63	0.27	0.16	0.05	0.28	0.59
Perceptual-Motor Coordination (CIRCUS No. 4)	0.44	0.53	0.50	1.00	0.47	0.52	0.32	0.13	0.17	0.32	0.51
Letter and Number Discrimination (CIRCUS No. 5)	0.49	0.62	0.47	0.47	1.00	0.49	0.47	0.31	0.15	0.43	0.52
Listening Comprehension (CIRCUS No. 9)	0.70	0.70	0.63	0.52	0.49	1.00	0.64	0.47	0.38	0.60	0.67
Productive Functional Language (CIRCUS No. 10)	0.51	0.44	0.27	0.32	0.47	0.64	1.00	0.47	0.42	0.98	0.36
Parts IA and IB	0.38	0.35	0.16	0.13	0.31	0.47	0.47	1.00	0.24	0.34	0.37
Quality	0.12	0.12	-0.05	0.17	0.15	0.38	0.42	0.26	1.00	0.25	0.24
Part II	0.50	0.43	0.28	0.32	0.43	0.60	0.98	0.34	0.25	1.00	0.32
Problem-solving (CIRCUS No. 13)	0.65	0.68	0.59	0.51	0.52	0.67	0.36	0.37	0.24	0.32	1.00
Age	0.24	0.25	0.13	0.24	0.17	0.20	-0.01	-0.02	-0.01	-0.01	0.13
Sex	0.05	0.08	0.10	0.17	0.03	0.08	-0.00	-0.07	0.03	-0.00	0.09
White	0.36	0.34	0.18	0.19	0.27	0.35	0.29	0.09	-0.00	0.31	0.08
Black	-0.28	-0.25	-0.21	-0.13	-0.21	-0.30	-0.21	-0.05	-0.05	-0.22	-0.09
Spanish	-0.16	-0.12	-0.02	-0.12	-0.09	-0.15	-0.18	-0.00	0.10	-0.22	0.03
Other	-0.10	-0.16	0.00	-0.13	-0.12	-0.05	-0.05	-0.14	-0.04	-0.02	-0.07
North Atlantic	0.14	0.09	0.11	-0.04	0.13	0.11	0.21	0.07	0.10	0.21	0.00
Great Lakes	0.01	0.02	-0.07	0.07	-0.02	-0.03	-0.12	-0.12	-0.15	-0.08	0.00
Southeast	-0.13	-0.11	-0.06	0.04	-0.06	-0.11	-0.10	-0.00	-0.03	-0.11	0.00
West/Southwest	-0.04	-0.01	0.02	-0.07	-0.05	0.00	0.00	0.06	0.09	-0.02	0.00
Professional	0.16	0.18	0.14	0.13	0.13	0.23	0.21	0.02	0.09	0.21	0.09
Skilled	0.10	0.07	0.08	-0.01	0.10	0.08	0.06	0.05	-0.08	0.07	0.03
Unskilled	-0.27	-0.26	-0.24	-0.13	-0.24	-0.32	-0.27	-0.12	0.00	-0.29	-0.13
Previous School	0.10	0.12	0.11	0.08	0.06	0.10	0.06	0.08	-0.01	0.05	0.13

## Appendix F

### SOCIAL AND PERSONAL DEVELOPMENT: INSTRUMENTATION

The discussion of social competence in Chapter 5 is conceptually organized, outcomes being derived from a set of theoretic notions about appropriate socioemotional influences of Head Start on children entering the public school system. Measures were treated in relation to outcome variables. This appendix is intended to supplement that discussion in two ways. First, it summarizes instrument recommendations, reorganizing them so that similar classes of instruments are discussed together. Second, it provides more detailed information about some of the recommendations. The appendix, then, comprises four subsections dealing with direct observation, evaluative responses from others, measures collected from subjects themselves, and measures restricted to subsample studies.

#### OBSERVATIONAL MEASURES

On the basis of a broad review of observational studies of children's social behavior, four stages in the process of instrument development have been singled out and summarized below. After general recommendations are drawn for the present study from these sources, more detailed procedural applications are suggested:

##### Steps in the Development of Observation Instruments:

1. The "target" behaviors, all and only those of interest to the study, need to be enumerated.
  - a. Exclusivity: some observation studies try to avoid multiple scoring of events (Ogilvie and Shapiro, 1969; Kimbrough, Barge, Bikson, and Smith, 1974), both for ease of scoring and determination of composite frequencies of event classes. Others (Bronson, 1973) multiply-code some events, allowing for the recording of dual-purpose interactions under each purpose category, and the like.
  - b. Exhaustivity: most studies do not require that all behaviors be scorable, social competence studies tending to focus only on interactive behavior. The Head Start Planned Variation report (Walker, 1973b) employs an observation

system scoring all child interactions, while Ogilvie and Shapiro (1969) have taken care to enumerate in their coding systems only those social behaviors "known to differentiate well-developed children from poorly-developed children." There is clearly some trade-off between precision and inclusiveness.

- c. For the present study, exclusive categories are recommended so that frequencies can be summed to create derived variables without devising complicated techniques to avoid counting the same event twice. Some arbitrariness necessarily results, since an event may occasionally exhibit properties relevant to two code categories; to avoid unreliabilities stemming from this source, priority rules should be built into the coding system (e.g., "if an event can be regarded as *child-initiated peer contact* and *child aggression*, it must be coded only as *aggression*"). Exhaustive categories, however, are not recommended. The present research is not interested in all child behavior styles but only those that bear importantly in a positive or negative way on social competence.
2. Operational definitions for target behaviors must be provided, so that scoring can be done without inferring attitudes, motives, traits, or dispositions. That is, characterizations of behavior must be as concretely descriptive and non-inferential as possible, so that persons familiar with the definitions could agree, on the basis of observation alone, how the behavior should be scored. (This requirement distinguishes observation techniques from behavioral judgments made after observation periods in which the observer judges whether, given what he has observed, a child should be called "active," "assertive," "cooperative," and so on.)

Because establishment of good operational definitions is difficult, and because instrument reliability hinges on this very point, it is recommended that pilot work only select and adapt elements from existing behavior-coding systems rather than attempt to generate any novel systems for the evaluation study. It must be emphasized that new code categories cannot simply be added without ascertaining their reliability and their importance for inferring social competence.

3. A scoring procedure must be devised, an area in which many technical decisions must be made.



- a. Choice of sampling unit: typically either time-sampling or event-sampling is attempted. In the tradition of behavior modification studies cited above, time-sampling is preferred. With a time-sampling procedure, score sheets are divided into time intervals of 10 to 20 seconds' length and each such interval receives a behavior score. This procedure was employed in the SDC Head Start data analysis (1972), where each child was observed for a total of 270 ten-second intervals; it is also recommended by Bronson (1973), who uses fifteen-second intervals. Other studies, such as the Head Start Planned Variation study (Walker, 1973b), the Ancona study (Jensen and Kohlberg, 1966), and other preschool assessments (Ogilvie and Shapiro, 1969; Grotberg, 1969; Kimbrough, Barge, Bikson, and Smith, 1974) employ event-sampling. Here a target behavior is recorded whenever it is performed by a child, so that frequency of scores represents frequency of discrete behaviors rather than time intervals.

Event-sampling rather than time-sampling is recommended because time-sampling procedures segment behavior into more refined units than seems necessary for the derivation of social competence indices.

- b. Spacing of sampling: where more than a single subject is involved in the observation study, it is necessary to decide the total time each subject should be observed and the way that time should be distributed. For example, the Jensen and Kohlberg study (1966) allocates one day of observation to every six children; with this procedure, the observer attends to three of the designated six children for five minutes and then the other three for the succeeding five-minute interval, alternating in this manner throughout the class day. In the study conducted by Kimbrough, Barge, Bikson, and Smith (1974), each of six subjects is observed singly by one observer in consecutive two-minute periods, the observer proceeding through the list three times in each observation session. A variety of other sampling systems are described in the observation studies reviewed by Walker (1973a). A sampling approach combining features of several such systems is suggested below.

4. After definitions and scoring procedures are established, inter-observer reliabilities must be ascertained. Typically such reliabilities are computed as the percentage of agreements to total agreements and disagreements in relation to the observation period as a whole and in relation to each category of behavior (the latter sort of calculation is useful for eliminating unreliable behavior categories or detecting category definitions

that need sharpening). Ordinarily, a sufficiently high reliability coefficient (0.80's) is sought before data collection is undertaken. If data collection points are distributed over a period of several weeks, it is desirable to repeat reliability checks at regular intervals.

- a. The research reviewed makes it clear that highly reliable observations can be made by nonprofessional observers after fairly short training periods (e.g., after less than a week of training with Bronson's quite complex time-sampling system, adequate reliabilities were obtained). The only exception to this general conclusion is found in Jensen and Kohlberg (1966), where observers attempted to score the free play behaviors of three children simultaneously. Overall reliability averaged only 47.8; but examination of the score sheets made it clear that the disagreement rate was boosted by one observer's missing an event that was seen by the other--a reflection of the fact that one observer cannot accurately track three preschoolers at once. When disagreements of that sort were eliminated, interobserver reliability reached 68.8 percent. Because such a rate was obtained when observers scored behaviors of three subjects at once, it is clear that observations on a single subject will pose no problem. Further, it is recommended that for indoor activity periods other than free play periods, observers should score behaviors of two subjects at a time.
- b. A second question turns on the amount of behavior time required to obtain a reliable estimate of subjects' peer, teacher, and task interaction styles. On the basis of preschool studies reviewed (e.g., SDC analysis of Head Start data, 1972, and the Jensen and Kohlberg analysis of Ancona data, 1966) it is concluded that 30 to 45 minutes of data per subject should be sufficient. It is further desirable that total observation time be organized to insure that adequate opportunities exist for observing a variety of subject behaviors in several contexts. This condition is implemented in the procedural suggestions below.

#### PROCEDURAL RECOMMENDATIONS

Three kinds of behavioral outcomes were suggested for measurement by means of observation instruments: child-child interactions and child-teacher interactions (both seen as subsets of behavior toward significant others in the public school), and child-task interactions (seen as an important though not primarily interpersonal academic role

behavior style). Although these outcomes are conceptually distinct, they should be assessed by the same observers during the same time in the public school year. Specifically, observations must be taken at a time reasonably coincident with the collection of peer and teacher evaluations (i.e., sociometric nominations and behavior ratings, respectively), in order that interaction data may be interpreted as supplementing these evaluations without history-maturation confounds. Within the basic observation period, however, data collection must be spaced to provide information about subjects' activities in different school settings. Peer interactions, for example, must be observed during free play periods, since other peer interactions during the school day tend to be directed by the teacher (Kimbrough, Barge, Bikson, and Smith, 1974). To observe a variety of teacher-child interactions, however, it is necessary that the teacher be relevantly involved in the ongoing activity (not just as a monitor, as during recess when the disciplinarian role is mainly involved) but that the activity not be overstructured (as during a test or flash-card answering period when all participation is highly controlled by the teacher). Finally, to observe child-task interactions, individual learning projects should be observed. For this purpose, it seems advisable either to devise a semi-standardized mastery task or to instruct the teacher in advance about the nature of the data sought; the latter course would at least insure a task focus, with the teacher not initiating any peer activities during that period.

It is recommended that pilot work on observing these three outcome classes be done by the same research team, whose aim would be to produce a single observation instrument with the three foci described above. A suggested organization of observation procedures follows.

1. There are three kinds of activities to be observed: individual learning projects (I); less formal semi-structured tasks such as arts or crafts (II); and free play (III). On the basis of literature reviewed, it is suggested that a total of 45 minutes of data be collected per subject, 15 minutes focused at each type of activity.

- a. Assume that all continuous observation periods are 20 minutes in length. (Longer continuous periods cannot safely be assumed, since in kindergarten and first grade classes attention to a single activity rarely lasts longer than 20 to 30 minutes.) Assume further that a single observer can simultaneously score *two* subjects during activities I and II, but can observe only *one* subject at a time during III.
  - b. Given that a typical class size for the present age group is 24, such a class in a sampled catchment area might be composed of one-third treatment subjects (T), one-third control subjects (C), and one-third who were not Head Start eligible or were not drawn for the study sample. Such a class would provide 16 sample subjects, eight in each condition.
2. For a class characterized as above, with 20-minute continuous observation periods, the observations should be segmented as follows; here A, B, C, D represent either pairs or individual subjects (depending on whether activity types I and II, or III, are intended) and numerals in parentheses represent time sequence:

20 CONTINUOUS MINUTES OF OBSERVED ACTIVITY

5 min: A		5 min: B		5 min: C		5 min: D	
2 1/2	2 1/2	2 1/2	2 1/2	2 1/2	2 1/2	2 1/2	2 1/2
(1)	(5)	(2)	(6)	(3)	(7)	(4)	(8)

Here pair (or individual) A is observed for the first and fifth 2-1/2-minute segment of the 20-minute periods; pair (or individual) B is observed for the second and sixth 2-1/2-minute segment; and so on.

3. On this system, three 20-minute periods will provide enough sampling time for activity I for the class, and an additional three 20-minute periods will provide enough sampling time for activity II for the class:

- a. T subjects: 4 pairs, observed in three 20-minute periods = 60 minutes @ two activity types (I and II) = 120 minutes.  
C subjects: 4 pairs, observed in three 20-minute periods = 60 minutes @ two activity types (I and II) = 120 minutes.
  - b. T + C subjects, for activities I + II = 240 minutes.
4. Because only one subject at a time can be observed for activity III, more total time must be focused here than on either of the other types; however, it is free play, an activity that is scheduled at least once and usually twice a day, so access to sufficient type III activity periods in a short span of time should not pose problems. On the system proposed here, six 20-minute periods will provide enough sampling time for activity III for the class.
  - a. T subjects: 8 individuals, observed in six 20-minute periods @ type III activity = 120 minutes.  
C subjects: 8 individuals, observed in six 20-minute periods @ type III activity = 120 minutes.
  - b. T + C subjects, for activity III = 240 minutes.
5. The total of all observation time for a class of 24 children, 16 of whom are sample children, is 480 minutes, or 8 hours (4 hours during directed activity of some sort and 4 hours of free play). The actual time involved will vary depending on the number of sample children per class, as will the time structuring. For example, if there are 16 sample children in a school number 16 divided between two kindergarten classes, observations may be structured into 10-minute periods per class (the observer moving between classes during task times but perhaps having access to all of them at recesses). Final scheduling will have to be done after the sample has been drawn. But given the basic condition that for part of the total observation time a subject is observed alone and for part of the time is observed along with another subject, the amount of time uniquely accounted for per subject is 30 minutes.
  - a. While T and C children were treated separately for ease of presentation above, it is in fact suggested that in any observation period both T and C children be included, mixing them as equally as possible.

- b. More than one 20-minute observation period may be obtained during a given day. While the same class probably will not repeat either a type I or type II activity, one of each might well be sampled on the same day along with a free play period. Order in which types of activities are observed is irrelevant.
6. Repeated observations are recommended for a subsample, after the entire sample has been observed in the early fall. Procedures would be the same as those established above, and time of data collection should coincide with the time at which other repeated measures are administered.

#### EVALUATIVE RESPONSE MEASURES

Many aspects of subjects' social competence were assessed by means of evaluative responses of other role incumbents. Because social competence turns in large part on being able to perceive role expectations and perform in ways that satisfy (or at least in ways that do not violate) them, responses from others are regarded as appropriate measures of subjects' social development. Such measures make use of existing instrumentation and are fairly easy and inexpensive to administer. What follows is an outline of all the evaluative response assessments recommended, and then a description of the Kelly role construct repertory test as modified for the present purposes; finally, archival information is presented in some detail.

#### Recommended Assessments:

1. Peers are the source of evaluative responses to subjects' interaction styles. Responses are to be collected early in the fall, at about the same time as process data are obtained through structured observations. It is suggested that observers be enlisted to collect this sociometric data just before beginning their observations, as a way of familiarizing themselves with the target population. Required for this data collection is a composite of photographs that includes everyone in the class; while recording of answers is very simple, some coding system should be worked out in advance so that answers would be ready-coded at the time of administration for keypunching directly from the response sheet.



- a. Three positive and three negative sociometric nominations are to be collected from every child in the class.
  - b. In addition, with sample children only, the picture-naming task should be administered.
  - c. It is recommended that the sociometric nominations be repeated at least once, in the second half of the school year, on a subsample basis.
2. Teachers are the source of five kinds of evaluative responses, listed below in order of time of administration.
- a. The earliest measures, to be taken within the first month of school, could be distributed through the school office in mimeographed form along with instructions and a self-coding response sheet. Teachers would be expected to complete and return such forms within the same day.
    - i. Scales of early adjustment are most suitably used during the second week of class.
    - ii. Summary estimates should be collected between the second week of class and the completion of classroom observations.
  - b. More extensive teacher ratings require someone (it need not be a professional) familiar with the rating techniques employed to administer the task. The task can be performed in a group, with as many teachers as can comfortably be seated at tables; a school cafeteria is recommended for this purpose. Both of the evaluation instruments listed here can be administered in the same session, with a brief break between them.
    - i. The Kelly role construct repertory test requires a Q-sorting tray, for sorting children in the class along teacher-generated evaluative dimensions. Here a heavy paper or cardboard fold out is suggested, with envelope-like pockets bearing scale values (cf. the Blocks' mailable Q-sort instrument). Index cards are needed on which to record names of constructs, along with sortable pieces of paper bearing children's names and subject identification numbers for ease of scoring.
    - ii. The Classroom Behavior Inventory will make use of the same Q-sorting tray. For this rating task, because the same standardized evaluative dimensions are used by all teachers, a duplicated set of scales can be administered. As in the previous task, however, sortable pieces of paper bearing children's names and subject identification numbers are needed.
    - iii. It is recommended that, for a subsample of subjects, these two ratings be collected again at least once during the school year.

- c. Social effect is assessed the year after the first public year, using the photograph composite originally devised for the sociometric task. Teachers are simply asked to name the children pictured, and indicate which were the best and worst students. If the schools are to be re-visited for any other purpose during the year after the major evaluation, the social effect measure could be collected at that time; if not, the assessment could be collected by mail (providing teachers with a self-addressed envelope, instructions, and response sheet).
3. Parents are the source of two sorts of evaluative responses. It is suggested that both be obtained during a single interview session at the home, perhaps at the same time as other family background information is being collected. These rating tasks should take little time since only one subject is being evaluated. It seems best that the interviewer read the items to the parent and record the response on a ready-coded answer sheet.
  - a. The Kelly role-construct repertory test will be administered first. No Q-sort materials are required for parents' use of this instrument. Some standardized probes should be devised in the event that the parents have difficulty thinking of constructs. The interviewer should be equipped with a set of rating scales whose anchor-points have been left blank, so that bipolar adjectives can be filled in as the parent generates them.
  - b. Summary estimates take less time, involving fewer items and relying on the sorts of judgments the parent is more accustomed to making.
  - c. The Kelly role-construct test is recommended for repeated measurement at least once during the school year with a subsample. Summary estimates are taken only once.
4. Observers are the final source of behavior ratings. These evaluative responses are to be collected in conjunction with the observer's regular behavior scoring duties; thus they involve very little additional expense while providing an extra insight into the subjects' performance relative to other institutional role incumbents.
  - a. The Classroom Behavior Inventory will be completed for each subject by the observer who scores his classroom behavior, immediately after the relevant observation



period. The form of the rating instrument should be the same as that used with parents, since observers will not have seen enough subjects individually to evaluate them by means of Q-sorting.

- i. Presumably observers can complete the rating task on their own if given an instruction sheet and if prepared for this task during their training sessions. The response sheet would then be returned to the research staff along with the observer's behavioral record for each subject assigned to him.
  - ii. If observations are collected on a repeated measures basis for a subsample, the same procedure should be followed with the Classroom Behavior Inventory.
  - b. The CIRCUS test number 16 focused at evaluating test-taking behavior will be completed by the person who administers the first cognitive test in the major battery. The rating response sheet will be returned to the research staff along with the test answer sheet.
5. Finally, archival information about the subject and his parents' school involvement should be collected from school records at the end of the first public school year. Although the data are obtained in a single visit, it should be organized temporally in the report forms so that performance during the first quarter of the school year can be compared with that during the second and third quarters. The main concern in preparing for the collection of archival information is deciding how to organize and record the data so they can be keypunched directly from the report forms and compared across school districts. (Additional discussion of archival information follows.)

#### DETAILED TREATMENT OF RESPONSE INSTRUMENTS

Besides the general instrumentation discussion above, three sets of response measures need further elaboration--namely the Kelly role construct repertory test for teachers and parents, and the classes of archival data just mentioned.

#### Kelly Role Construct Test as Modified for Teachers:

1. The teacher is asked to think of students whom she has taught and is now teaching, and then make notes of: (a) the best

student she ever had; (b) the best student she had last year; (c) the next best student she had last year; (d) the worst student she ever had; (e) the worst student she had last year; (f) the next worst.

- a. If any of these overlap--e.g., (a) and (b), or (d) and (e) might--the teacher is asked to substitute another relevant name.
  - b. The teacher is told she will not be asked to tell the interviewer who these children are; they will just serve as reference points in answering the remaining questions.
2. The teacher will then be asked to make a series of comparisons of the following sort: "Look at the first two names you have on your list, (a) and (b). Think of one characteristic they have in common that (d) does not have. [pause] On the first index card in your deck, please write down that characteristic and its opposite."
- a. Comparisons and contrasts proceed in this fashion until the desired number of bipolar dimensions have been elicited, typically between 10 and 20. Piloting would be useful to determine the best way to modify Kelley's instructions for the present project. (On the basis of experience with this test we think that 10 characteristics should be sufficient, and that asking for more than 15 would produce fairly redundant constructs toward the end.)
  - b. What results from this part of the test is a deck of index cards for each teacher, each card representing an evaluative dimension. Moreover, these evaluative dimensions define for that teacher, on the basis of her own experience, the salient role-qualities of students. At this point, the list of names used to generate the constructs may be discarded.
3. The second part of the test is a simple scaling task, where the index cards form the scales and the students in the class are the subjects to be rated by means of the scales. (At this point, each teacher is using the same number of scales, although the content of scales will differ somewhat from teacher to teacher.) A seven-point scale is recommended so that scale-values will be comparable to those elicited using other

standardized rating scales. Further, it is recommended that rating be performed by means of a constrained Q-sort, to overcome the distribution differences that otherwise might arise (creating problems in between-rater comparisons).

To perform the rating task, each teacher must be supplied with a deck of cards bearing the names of students in her class, along with a sorting tray whose compartments bear the scale designations (extremely, moderately, slightly, neutral, and so on) in the proper order (Block and Block, 1972, Q-sort materials).

- a. The teacher is asked to turn over the first index card, which bears the first bipolar scale. The children currently in the class are sorted, in a flattened-normal distribution, into rating categories.
  - b. The same procedure is then followed for each of the remaining construct dimensions. At the end of the rating task, each child has the equivalent of a score (range 1 to 7) on each of the rating dimensions. It is recommended that the positively keyed end of the scale be consistently represented by the higher numerical values so that raw scores will simply represent degree of favorableness of evaluation.
4. The most important information provided by this evaluation is the extent to which children satisfactorily enact teachers' role expectations, and the degree to which (if at all) Head Start children differ from control children in this very important respect.
- a. While it is expected (given task instructions to teachers) that the constructs all represent a general evaluative judgment, it is worth determining by factor analysis whether this judgment is uni- or multi-dimensional. That is, some constructs generated by teachers may be social in nature and orthogonal to task-related constructs (as has been found when scales are supplied to teachers rather than generated by the teachers); however, it could be that teachers' own evaluative judgment domains are undifferentiated and one-dimensional. In any case, it is important to know the extent to which standardized rating scales represent the teachers' ongoing evaluative responses with respect to item content and response structure.

- b. It should also be interesting to examine degree of overlap in terms defining student role qualities for teachers of the same school system, the same national region, the same race or status background, and so on.
  - Administering this sort of teacher evaluation during the Head Start year to a subsample would further allow determination of the extent of overlap between Head Start teachers' expectations for their students and those of public school teachers.
  - Similarly, use of the role construct test will permit comparing teachers' expectations with those of parents.

Kelly Role Construct Test as Modified for Parents:

1. The parent is asked to think of examples of children who are or were very good or very bad in school. The purpose of these examples is to help the respondent generate concrete, behaviorally meaningful characteristics. No given number of examples is required, although it is important to have both positive and negative role models. The following probes might be useful for pinpointing such instances: (a) a sibling who is or was very good in school; (b) a neighbor's or relative's child who is or was very good in school; (c) the ideal school child, as the respondent imagines him or her; (d) a sibling who is or was very bad in school; and so on. It is important to remind the respondent that these names need not be revealed.
2. The parent will be asked to make a series of comparisons and contrasts, in order to analyze salient role qualities of school children from the respondent's viewpoint.
  - a. It is important to obtain not just one adjective but also its polar opposite in order to form a rating dimension.
  - b. In the one-to-one parent situation, the interviewer should write the constructs down as the respondent decides on them. Because the parent will be rating one child rather than sorting a group of children, regular seven-point scales can be prepared in advance for the desired number of construct dimensions. The interviewer can then enter each set of contrary characterizations over a rating scale.
3. The second part of the task involves asking the parent to rate his own child on each of the construct dimensions just generated, using a typical seven-point scale where higher scale values are given to positively keyed items.

4. The most important information provided by this evaluation is the extent to which children satisfactorily enact parents' role expectations, and the degree to which (if at all) Head Start children differ from control children in this respect. However, response-biasing can be expected, parents being reluctant to give bad evaluations of their children to an unfamiliar interviewer. To the extent that response-biasing occurs, it will minimize differences between Head Start and control subjects.
  - a. A second important sort of information to be gathered from the parent role construct test is the degree of overlap between the parent and the teacher views of salient role qualities of school children. Such information would indicate whether "social competence" means similar things to parents and teachers.
  - b. Degree of congruence among parents' expectations for parents of children in one school system should be established, in order to determine whether a common set of role expectations exists and whether expectations differentiate parents of Head Start children from parents of control children. Should such differences exist, it would be additionally interesting to see whether either set of expectations was significantly more similar to teacher expectations.
  - c. The set of parent-generated constructs is to form the basis for judgments concerning size of value-discrepancies between the child's family or ethnic group and the school as a representative majority culture institution. It will thus aid in the interpretation of measures of attitudes and behavioral styles (Chapter 5, parts 3 and 4).

Archival Information on Students' Performance:

1. Placement, tracking, or "special class" information should be obtained from all school systems where incoming children are assigned to classes based on some sort of performance criterion. Walker (1973b) points out that being assigned to a "slow" or "normal" or "bright" class has long term implications because of its joint effect on teacher expectations and children's academic self-images (and their interaction). It is well known (Gerard and Miller, 1971; Coleman, 1966) that lower-status and minority children are placed with disproportionate frequency in retarded classes. To the extent that Head Start makes a difference in school-task and school-test performance (as well as

in general ability to comply with student role expectations), it may affect placement; and the long term importance of initial placement--given the self-perpetuating nature of such diagnoses--is inestimable.

2. School attendance and tardiness records should be available from all schools participating in the study. These measures are recommended by Cowen et al. (1965) and Kohlberg, LaCrosse, and Ricks (1972) as symptomatic of school adjustment. In particular, truancy (or unexcused absence) is regarded as an important index by Kohlberg, LaCrosse, and Ricks.
3. Records of referrals to the school nurse's office ought to be obtainable from all participating schools and are suggested for collection by Cowen et al. (1965), who found visits to the school nurse one of the most sensitive behavioral records in their study of third-graders. They found it exhibited a significant relationship to Children's Manifest Anxiety Scale scores and was associated with teacher ratings of maladjustment as well as self-ideal discrepancy scores.
4. School success or failure indicators:
  - a. School success indices may include the following items.
    - Nonacademic grade reports (e.g., "courtesy," "good citizenship"), as well as teacher comments on the report or in the child's school record;
    - Official "goodness" recognition practices, which vary from school to school and include such things as being a "class monitor," receiving "good citizen" awards, being elected to a class office, and the like (recognition practices would have to be ascertained from each school);
    - Positive integration of the school into the child's life, as indicated by use made of the library, checking out school play equipment after hours, and other extracurricular opportunities used; such "extra-curricular" or voluntary occasions for use of school facilities will have to be determined for each school.
  - b. School failure indices may include the following items:
    - Nonacademic grade reports (e.g., "talks out of turn," "interrupts classroom activities," etc.) as well as teacher comments on the report or in the school record on the child;



- Official "trouble" indicators, which may vary from school to school but typically include visits to the principal's office, being sent home, or less severe punishments such as suspension of recess or isolation during a group activity;
  - Strongly antisocial behavior indicating inability to accept the school situation, including official constraints placed on the child by the school, recorded incidences of vandalism, and violent or aggressive behavior infringing on personal rights or property of others (Kohlberg, Lacrosse, and Ricks, 1972); according to Kohlberg, Lacrosse, and Ricks, the single best long-term indicator of social adjustment is frequent or severe antisocial behavior (where antisocial behavior primarily refers to violation of implicit or explicit rules about either personal integrity or maintenance of group cohesion and interpersonal trust).
- c. For purposes of aggregation of data, it is recommended that exploratory work be done aimed at devising a comparable coding of all data bearing on school success and failure. Two options are available. Most simply, each child could be assigned a single scale value based on a review of his records as outlined in a. and b. above; interjudge reliability could probably be obtained after some criteria had been established, so that scale values would be global indices of school role success. Alternatively, some way of classifying recorded items from important to unimportant instances of success or failure could be established; once interjudge reliability had been ascertained, items could simply be counted and each child given a frequency score per item class.
5. Pilot investigation should determine the best indices of parent involvement and how to code them for purposes of aggregation. Presumably these variables will be either the same as the corresponding parent background variables derived for the Head Start year or else a subset of these (should some of the Head Start parent involvement categories be inapplicable in relation to the public school setting).

#### SUBJECT MEASURES

Subject measures are those collected from subjects by an examiner or interviewer outside the classroom, either on a group or on an individual basis. While the administration of subject measures does not require a professional psychologist, some training would be required to familiarize the examiner with procedures. In many cases, pilot

research is necessary before a final recommendation can be made concerning the manner in which an outcome should be assessed. Typically, problems focus on selecting among available stimulus materials or procedures rather than on interpreting or scoring behavior of subjects. Similarly, most time and expense invested in subject measures is related to establishing a standardized situation in which behaviors may be elicited rather than to observing and coding those behaviors. For the measures presented below, time of administration is not particularly of concern although it would be preferable that they *not* be used during the first few months of the school; it is assumed that these instruments could be combined with instruments from other evaluation areas and administered at a time determined by convenience in relation to overall test scheduling. None of the assessments below are recommended for repetition within the first public school year, but a few of them are suggested for administration during the Head Start year on a subsample basis to provide developmental data. Primarily, however, the measures are to be administered once only, sometime after the Christmas vacation, during the first public school year. Three sorts of information are provided in the outlines of subject measures below: First, a summary of recommended measures is given; next, two closely related measures requiring considerable pilot development are discussed; third, a set of individual performance measures of outcomes proposed for evaluation not ultimately included among recommended subject measures is discussed, along with reasons why they were excluded from the noncognitive battery.

Recommended Assessments:

1. Self-report or preference data are the target of the first major class of subject instruments. These instruments are typically easy to administer and should be inexpensive once selection of stimulus materials and piloting of procedures have been completed. It is assumed that some of them are to be individually administered, but pilot investigation should always include exploration of using them on a small group basis instead. In every case, preliminary work should include the devising of a self-coding response form to be completed by the examiner (so that data can be transferred



directly into a computer). Instruments of the self-report or preference type are listed below in order of presentation in the text of Chapter 5; they are intended for administration to the entire subject sample, unless conclusions drawn on the basis of pilot effort render such administration infeasible.

- a. All three role-taking instruments are of the same sort, involving responses of subjects to questions about picture stimuli. It is envisioned that the three instruments will be subparts of the same examination booklet. Aspects of role-taking are ordered in terms of the complexity of social comprehension they involve. Although it seems that the role-taking assessment would have to be conducted individually, pilot research might consider ways of operationalizing it for a group.
  - i. Spatial perception is a Piagetian task focused at egocentrism-sociocentrism. Pilot work should select an existing, brief measure.
  - ii. Situational perspective involves the discrimination of roles within the family and school role system, based on Emmerich's research (1959). Pilot research is required to develop stimulus materials in addition to the family-related stimuli used by Emmerich.
  - iii. Cultural perspective requires the discrimination of proprietary norms or majority culture values based on Scott's work in Australia (1969). Preliminary work must develop all stimulus materials, based on Scott's examples but with content relevant to this culture's socialization emphases.
  - iv. All these role-taking measures except the school-related stimuli in situation perception are recommended for subsample measurement during the Head Start year with treatment and control subjects.
- b. Response range, an outcome area related to alternative role enactments in response to interpersonal stimuli, is to be assessed in two ways (individually).
  - i. As-if or consequential reasoning has been suggested for assessment using photographs of teacher-like faces in varied affective states; after context descriptions, children would be asked what their teachers would probably do in response to given child behaviors. Such a measure would have to be developed completely in pilot research. Should such a task prove unfeasible, the Spivak and Shure "What Happens Next?" game is recommended (1974).
  - ii. Responses to interpersonal conflict situations are measured by the Spivak and Shure PIPS test. This measure needs no pilot work except for the following

decisions: Which half of the test should be used; in the event that a task of consequential reasoning about teacher-behavior cannot be developed, is the "What Happens Next?" game sufficiently independent of the PIPS test that administering both measures would be worthwhile? If the latter question is answered negatively, only the PIPS should be used; if affirmatively, no more than half the "What Happens Next?" game should be used.

- c. School attitudes are to be evaluated by means of a verbal self-report based on PASS and the Minuchin et al. (1969) sentence completion test, including only those items that pilot exploration decides are important indices of subjects' feelings about their role in the academic setting. It is suggested that the alligator game (McNeill, 1970) be used as the question-answer medium, probably with the sentence-completion format. The measure will presumably be individually administered.
  - d. Attitude toward intellectual challenges and accomplishments is measured by a picture selection task, the Children's Achievement Wishes Test (Crandall et al., 1962, 1965). Pilot investigation should be undertaken to decide whether to use all or some of the stimulus pictures, and whether it can be group administered.
  - e. Self-attitudes are measured using the Ziller Self-Social Constructs Test (described in Walker, 1973a). Two items are definitely intended to be included, the self-esteem measure and the distance-from-the-teacher measure. Pilot researchers might wish to consider the inclusion of other items as well; if other items are included, a decision needs to be made whether it is worthwhile to order the response forms or produce them. The measure is administered individually.
2. The second set of subject measures involves the elicitation and scoring of subjects' behaviors in a variety of standardized situations. Such assessments are recommended because it is assumed unwise to rest the entire evaluation of social competence on verbal reports from subjects and others who react to their behavior, and because not all important aspects of behavior involved in social competence can easily be investigated during naturalistic observation sessions. The following instruments involve considerable preliminary work on procedures. Scoring is expected to pose no problem, and it should be easy to render responses in self-coding form.

- a. Direction-following and task-completion are grouped under learning styles and recommended for assessment in two ways, the simplest of which is to be included in the test battery. It is suggested that pilot work first devise a structured mastery task (cf. Bronson, 1973) to be performed by subjects during a regular class period; their work would be graded later. Otherwise, such a task could be group administered separately and graded after it was completed. If neither of these alternatives is satisfactory, the Blocks' dual focus task (1972), already fully developed, is recommended for individual administration.
- b. Goal-setting and self-evaluation need considerable pilot work but are nevertheless very important to include in the entire battery. Preliminary research should be based on the work by Crandall and others (1962, 1965) and Weiner (1972). It should aim at combining these resources to produce a single complex task involving the following steps (individually administered):
  - i. Selection of task difficulty level (here difficulty levels would be described only as "hard for children your age," "very easy....," or "about right/medium...." Although the subject's response is scored for level selected, in fact all subjects will be presented with exactly the same task).
  - ii. Designation of minimum achievement level (here the subject is asked to point to the task so easy he would be mad if he couldn't do it because "even a baby could do it"); a graded series of tasks, preferably mazes but perhaps puzzles, must be devised with pilot work determining how many should be in the series.
  - iii. Designation of success expectancy range (here the subject is asked to point out the other items in the series he thinks he can successfully complete); pilot work must establish that some items are sufficiently hard and some sufficiently easy to make this indication interesting.
  - iv. Actual performance of the series of tasks; here subjects are asked to do as many as they can.
  - v. Self-rewarding behavior; subjects are told they may reward themselves as they see fit from a supply of prizes. Pilot work here must decide what to score (of course rewards will be scored in relation to difficulty level and success of outcome; the scorer should probably also note level of reward in relation to amount of effort vested in each task).
  - vi. Attribution test; this is a verbal step probing the subject's self-reward style (does he think rewards should be dispensed on the basis of effort or ability or outcome?). Questions are posed in terms of an earlier (imaginary) subject who could not decide how to reward himself.

- c. Resiliency in response to a frustrating nonpersonal situation will be measured unobtrusively using the Blocks' (1972) "stuck drawer" technique. Procedures and scoring are fully worked out.

Role-Taking Measures to be Developed:

1. A discussion of Emmerich's theoretical assumptions involved in the development of his family role-discrimination test is presented in the text of Chapter 5. Emmerich's (1959) initial assumptions should be retained and brought to bear on the school setting. Clearly the power dimension should be retained for the school social system, but this dimension can be further articulated by considering what aspects of power differentiate parents from teachers in relation to the child, and what sorts of expectation differences the child bears toward the use of such power in his role as student versus in his role as family member. The French and Raven (1959) social power taxonomy might be a fruitful starting point for these distinctions. Further distinctions might turn on the context in which power is relevantly used (or kind of power). Similar remarks bear on the function dimension, which should distinguish parent from teacher roles in relation to both kinds of expressive functions and kinds of instrumental functions. For teacher functions, the teacher-child interaction categories in Bronson (1973) and Ogilvie and Shapiro (1969) might provide starting points. Additional dimensions may be incorporated in the conceptual minisystem.
2. A set of stylized figures should be developed to represent role incumbents. It is recommended that stimulus figures include both family and school roles, since it is important to determine whether the child is learning to distinguish role behaviors relevant to the two social contexts. Thus when the child figure was paired with the mother figure, family role behaviors would be involved but when the child figure was paired with the teacher figure, school role behaviors would be involved.

- a. Pilot work will need to determine how many figures are needed. It might be possible to eliminate the father figure, retaining only the mother and teacher as adult figures. Further, it might be desirable to eliminate one or both siblings in favor of school peers, or else to leave the child figures unidentified except as children. (Because sex of child figure yields some role distinctions not relevant here, perhaps the best course would be to use only figures of the same sex as the subject.)
  - b. Pilot work also must involve the development of test sentences that adequately convey role relationships without exceeding the comprehension of the subjects.
    - i. The most popular role-taking tests avoid this difficulty by asking the child himself to generate descriptions of roles or to tell who says what to whom (the Racial Attitude Sex-Role Picture Series, the Social Episodes Test, and the Social Perception Interview reviewed in Walker, 1973a).
    - ii. However, Mussen's (1960) review of such tests indicates that responses often represent recent real events rather than role perceptions, and that scoring is extremely time-consuming. He concludes that if one desires the child to attend and respond to a specific and well-defined topic, it is best to use well-structured pictures along with interview techniques that limit the range of responses.
  - c. Emmerich's (1959) general response elicitation procedures are therefore regarded as most satisfactory and should serve as models for generating additional school-situation test sentences.
3. Scott's (1969) test of perception of proprietary norms or cultural expectations is, to our knowledge, unique in content. The method of test construction as well as the test procedures should follow Scott's as closely as possible. The basic research problem is very much like that represented by Emmerich's work and should be adapted for the present purposes in much the same way. That is, the initial assumptions with which Scott begins should be retained and brought to bear on the present problem. Specifically, it is assumed that there is a strong set of cultural values implicit in the majority culture and explicit, perhaps, in the OCD-Head Start Policy Manual (1973) concerning what characteristics are important to develop in children as a part of their socialization. A small set (two to four) of values should be identified as having priority.

After key values are identified, teachers and mothers must be questioned to determine what cultural expectations exist about how children should behave in situations where these values are relevant.

- a. For this purpose, the researchers must devise a set of relevant situations representative of events that could naturally arise in the child's home, neighborhood, or school environment about which to raise questions.
  - b. The cultural expectation is equated with the modal response from each ethnic group. While perception of majority culture values (represented by teachers' modal response) will be of special interest as the primary outcome measure, cultural expectations bearing on the same situation held by different demographic groups will help interpret subjects' responses. It is assumed that Head Start will produce most significant effects where the discrepancy in cultural expectations is greatest.
4. After this conceptual part of the instrument development is completed, results must be translated into stimulus materials and test questions for children.
- a. It is suggested that some imaginary role incumbent be chosen (either from the stick-figure or animal world) as the identification figure for the child. Pictures must be generated that involve this figure in the situations about which perception of cultural expectations are to be tested.
  - b. Test questions should be devised to go with the cards; Scott's method of using only questions that can be given "yes" or "no" answers should be used. In any case, guidelines given above for response elicitation should be followed here.

#### ELIMINATED MEASURES OF LEARNING-RELATED ATTRIBUTES

There are a number of approaches for eliciting and assessing behaviors thought to be related to school success. Some measures were reviewed for present purposes and ultimately excluded from the final battery (in favor of less familiar measures, in some instances).

Reasons for these exclusions are presented below.

1. Impulsivity is regarded by many sources as an important control mechanism reflected in response tempo and indicating an inability to attend to the problem, deliberate, and then respond accurately. This characteristic is typically measured



either by a motor inhibition test (in which response duration is timed after a child is instructed to do something as slowly as he can) or by response latency on the Matching Familiar Figures test (where short latencies are assumed to interfere with otherwise attainable accuracy).

- a. Motor inhibition is not obviously related to the kind of academic tasks children typically face in school, and it is not at all clear that delaying of responses is a cross-situational trait (Kohlberg, Lacrosse, and Ricks, 1972). Nor is it clearly developmental; Ward's (1973) analysis of Head Start longitudinal data yields only a slight increase in latency with age (from four to six years), with sex and SES differences being quite erratic over time. Moreover, this measure correlates more closely with general ability measures than with other response tempo measures. Further, the groundwork done with this measure in field tests by Boger and Knight (1969) yielded data so skewed as to be unusable.
  - b. The Matching Familiar Figures test has also been used to collect latency data. Ward's (1973) longitudinal data on this measure yielded little of interest (males tended to be slower and also less accurate than females) and was more closely associated with general ability than with motor inhibition results. Emmerich (1973), however, found no association between latency and skill level. Perhaps the most definitive critique of the Matching Familiar Figures latency scores as an impulsivity measure comes from the Blocks' analysis of their own results with this test (1973). On the basis of their work it seems clear that any such latency measure would be useless as a test of reflectivity-impulsivity.
  - c. While Shipman (1973) finds response tempo consistently emerging as a second strong factor in the longitudinal analysis orthogonal to general ability, such a factor is evidently not measured by either of the tests cited. To the extent that it is an influence on task or test performance, then, it will have to be reflected in the "deliberation" category of the Bronson (1973) executive skill coding as recommended above and in the tempo ratings included among the factors affecting test performance.
2. Persistence might be regarded as a behavioral characteristic reflecting low-level resiliency--i.e., it represents the subject's willingness to undertake the same task anew or continue it, relative to some purpose. While many sources (Boger and Knight, 1969) have measured persistence as if it were a virtue, it is not far removed from typical measures of rigidity

(Shallenberger and Zigler, 1961; Zigler and Butterfield, 1966). That is, rigidity is often measured by a subject's willingness to repeat the same response over and over, as exemplified in standard "cosatiation" tasks. Low-SES subjects are often found to be more rigid on such measures than their higher status counterparts, according to Zigler, because of their desire to please the experimenter. Persistence research, however, usually concludes that lower status subjects are less willing to persist. (It is seldom interpreted that lower-SES subjects are more persistent given satiation tasks or less rigid given persistence tasks, even though the approaches differ hardly at all.) Because of the possible confound between persistence and rigidity, neither construct is recommended here to be directly measured (the two are clearly related and seem to fall very near the middle of the flexibility-rigidity axis).

3. Delay of gratification is also frequently regarded as an important control mechanism, typically measured by arranging a situation in which a child is to receive a reward; the child is then permitted to choose between a larger future reward or a smaller immediate one (usually the reward is some kind of candy). Ward (1973) found in his analyses of Head Start data that the delay of gratification measure showed no effects for any independent variables and correlated with nothing else. Boger and Knight (1969) also found the measure unrelated to independent variables and uncorrelated with other dependent measures from the Cincinnati Autonomy Test Battery. Finally, the Miller and Dreger literature review casts considerable doubt on this measure even when it does show between-group effects (1973).
4. Field dependence is often recommended (Baumrind, 1974) as representative of inner-directedness or self reliance. It is typically measured by the Witkin rod-and-frame task (Block and Block, 1973; Gerard and Miller, 1971). This task is scored with respect to dependence on external field or background cues for



perceptual judgments, where such dependence is negatively related to accuracy. The results of large scale analyses on this measure are unsatisfactory (Gerard and Miller, 1971), and a thorough literature review casts doubt on operations of this sort as measures of inner-directedness in tasks at all (cf. Rotter, 1960).

5. Both the Murray Thematic Apperception Test and the Bellak and Bellak Children's Apperception Test have been used to measure achievement orientation. In both cases, the subject is shown a series of standardized pictures designed to elicit achievement themes; for each such picture, the child is asked to compose a story. As projective techniques, both tests are time-consuming and costly to score: The entire story must be recorded, transcribed, and coded in respect to many aspects of the achievement orientation. In addition to this difficulty, however, another makes these measures infeasible even for small samples: In relation to both tests, it has been found that young children tell such short and simple stories that either they do not contain achievement imagery or else the achievement themes are not scorable in terms of the complex scoring procedures available for them (Crandall et al., 1962; Gerard and Miller, 1971).
6. The Gumpgookies test is a semi-projective picture interview designed to measure school achievement motivation. The test includes 27 forced-choice items in which the child is asked to select the gumpgookie "most like me." This test is recommended by Shipman (1973) and Stearns (1971). However, Walker's review of the measure (1973a, b) as used in a large Head Start sample as well as other populations yields the following conclusions. First, the 27-item form is too long for preschoolers, but attempted shorter forms have unclear reliabilities (presumably because the original 27 items were needed to measure the five components of the Adkins and Ballif achievement model). Second, the items themselves have a clear social desirability bias, which may account in part for the obtained reliabilities and

is responsible for ceiling effects. Finally, the test does not correlate with achievement, although it does correlate with teacher ratings of achievement; these considerations suggest that the test, like teacher ratings, tends to be sensitive to middle class response style rather than to achievement behavior.

7. Behavioral tasks such as ring-tossing, bean-bag-throwing, and dart-throwing have been employed as less obtrusive indices of goal-setting and self-evaluation (Gerard and Miller, 1971). In such cases the child chooses how close to the target he wishes to stand, and guesses how many times (out of ten trials) he will hit it (McClelland, 1953, after which most of these tasks are patterned). The child is offered one or more opportunities to adjust his distance and estimated success frequency. This approach to achievement behavior is regarded as basically preferable to the two discussed above, but the class of behavior used in the goal-setting task is not representative of school-related achievement situations. Second, the task tends to discriminate on the basis of physical dexterity and past practice with similar games (e.g., Gerard and Miller found that minority males had the most conservative and also extremely accurate estimates of where they could stand to be 100 percent successful). The Block tower task is susceptible to the same sort of criticism (Block and Block, 1973), except that most preschoolers have had experience with some sort of construction tasks.
8. Locus of control has been regarded as an important attitude variable related to academic and social competence, the concept of self as agent being the attitudinal variable most closely associated with achievement in minority and lower-status populations (Coleman et al., 1966; Battle and Rotter, 1963). Recent reviews of locus-of-control literature suggest, however, that most such tests are culturally biased and fail to distinguish different domains in which the agency of the self is experienced. For example, both Gurin et al. (1969) and Miller and Dreger (1973) contend it is minimally necessary to distinguish internal control as personal efficacy from

internal control viewed as responsiveness of the social system to individual efforts; these sources find Black subjects scoring fairly high in the former but not in the latter domain, a difference easily interpretable given the discriminatory social system within which Black subjects must operate (Williams, 1970a). Instruments that do not make this distinction typically find Black respondents globally low in internal control, too general a conclusion. And, for the present purpose, it is necessary to restrict the range of controlled consequences to academic ones.

Although most locus-of-control instruments are applicable only to older populations, exploratory work could generate an appropriately restricted instrument usable for subjects in the present age range. A review of relevant literature does not, however, offer promising candidates for age modification. Thus we think the best course is to examine achievement behavior in an experimentally controlled situation and question subjects about their performance. Locus-of-control measures per se are recommended for exploration in focused study.

#### SUBSAMPLE MEASURES AND FOCUSED STUDIES

Other instruments measure important outcomes but are not included in the entire battery either because of the time and effort involved in their administration or because of uncertainties regarding how adequately they could be developed in the allotted preparatory time. Such assessments are reserved, in important cases, for either subsample evaluation or focused studies. Typically the distinction between what is feasible for a subsample and what must be recommended for a focused study has turned on how exploratory the proposal is; in general, subsample studies are less exploratory and provide systematic supplementation to results obtained for the entire sample; focused studies are more investigative, require or may even simply be pilot research, and will help interpret aspects of social competence development. What follows, then, is first an outline of outcomes proposed for subsample measurement or focused study and last a procedural supplement for some of the subsample evaluations.

Proposed Assessments:

1. Subsample measurements include some observation and rating instruments but primarily involve subject assessments. For these evaluations, time of administration is not important and should be determined by convenience.
  - a. An unobtrusive measure of children's interactions during indoor free play periods is recommended using time-sampling by means of cameras taking photographs of the behavior space at a set frequency. Pictures would be scored to indicate frequency of contact with other children, group size, and ethnicity of play associates (as well as frequency of isolated play); identity of play contacts could also be recorded by scorers familiar with the children, to corroborate sociometric data as well as observations of peer interactions. This technique, should it be successful, is expected to provide an inexpensive and reliable alternative to structured observation.
  - b. Interpretation of evaluative constructs used by teachers and parents is to be enhanced by a study of the connotative meanings and response structures surrounding these constructs.
    - i. Teacher responses to the Classroom Behavior Inventory are to be investigated by using the long form (the major item pool) from which it was generated. There are several large item pools available, including the California Child Q Set, for this purpose. Research efforts here involve insuring that there are enough rated children in each major ethnic group to perform required statistical analyses, and more important, performing an appropriate set of analyses. Factorial replication of factor structures within ethnically defined groups and multiple discriminant analysis are among the analytic procedures recommended. Actual data collection is unproblematic since rating scales and procedures are fully developed.
    - ii. A semantic differential study is suggested to explore use of evaluative terms among teachers and parents in characterizing subjects of different ethnic groups.
      - Pilot steps are needed to decide exactly what the stimuli should be; at minimum, pictures of children easily identifiable as representative of different subcultures should be devised, and the usefulness of combining these with auditory stimuli should be considered.
      - The number and content scales should be readily decidable after brief preliminary testing of existing semantic differential instruments.

- The number of respondents needed, and their introduction to the task, has to be decided. Actual data collection should not pose any difficulties. Of central interest is the extent to which the same constructs load on different factors as a function of the subculture of the rater and of the subject being rated.
  - Some method of estimating distance between subcultures on these parameters should result.
- c. Subject learning style measures for subsample study:
- i. A measure of learning of intentional versus incidental responses has been recommended for subsample study. Approaches to be used as a basis for evaluating this learning style are reviewed below.
  - ii. A measure of reinforcement style is recommended for subsample study, attempting to see whether Head Start children are more responsive to correctness feedback than are control children. Experimental techniques for this assessment are discussed below.
  - iii. Epistemic motivation, a subset of curiosity that is specifically relevant to academic goals, is a learning style outcome and is recommended for measurement by a picture choice task. Preliminary investigation of the work done by Maw and Maw (1962) with this instrument is needed to determine whether to keep all or some of the original stimulus pictures or devise new ones for inclusion. The number of choices to be made is also to be determined. It is a candidate for group administration.
- d. Among the subject measures of role-taking and response range, one is recommended for subsample study. Among existing ways of assessing multiple correct responses to an unstructured stimulus situation described in the procedural supplement below, one should be selected to evaluate resiliency in response to a nonfrustrating, nonpersonal situation.
- e. Attitudes are to be evaluated in a subsample interview whose aim is to provide criterion values for variables indexing the subject's feelings toward himself and the academic setting. Interview procedures are discussed more fully below.
2. Focused studies will involve extended research. Questions such as time, place, and manner of administration are not relevant here. Rather, the list below indicates only what areas will be of interest.
- a. The first class of focused studies deals with the development of behavioral-experimental assessments of characteristics regarded as important to the notion of social

competence and for which existing evaluative techniques are not regarded as wholly satisfactory. The characteristics targeted here include:

- i. Responses to interpersonal conflict situations.
  - ii. Locus of control.
  - iii. Attitudes toward school (picture-taking interview, videotape coding).
  - iv. Self-and-school role congruence (attraction-similarity approach; World Test).
- b. Multiple role integration (exploratory study aimed at determining the nature and number of role minisystems in which subjects participate, the demands and evaluations they make, and adaptive and maladaptive ways of coping with multiple roles).

**PROCEDURAL SUPPLEMENT TO SUBJECT MEASURES RECOMMENDED  
FOR SUBSAMPLE STUDIES**

In some instances, outcomes suggested for assessment in subsample studies above require pilot research directed at the development of an appropriate measure from existing techniques. The procedural supplements for these outcomes (provided below) represent assessment approaches that seem most promising for such studies after extensive literature reviewing.

1. Learning of intentional and incidental responses has been investigated by researchers interested in imitation and modeling, as well as by researchers concerned primarily with influences on school success. Two experimental paradigms are reviewed here which are seen as most promising for use in evaluating this aspect of learning style.
  - a. The Postman Game devised by Ross (1966) is an individually administered task that children seem to enjoy very much. The experimental variables include two sorts of responses, intentionally instructed and incidentally learned (an equal number of each are involved). The manipulation is delivered in an area contrived to represent a post office, including counter, telephone, mail slots, play stamps, and money drawer. The examiner tells the child she will teach the child how to be the postman. Instructions include, for example, "When the telephone rings, do not say 'hello'; say 'post office.' Remember, say 'post office' instead of 'hello.'" But when the examiner actually



answers the telephone in that way, she also puts her foot up on the chair in an idiosyncratic manner. For each instruction there is an accompanying incidental behavior.

After the instruction session, the child is allowed to be the postman in the Postman Game. The telephone is rung by the examiner, and children not serving as subjects come in to mail letters. Learning of intentional and incidental responses is scored, and then the child is asked to teach a new subject how to be postman; here instruction of intentional responses is scored.

- b. The Geography Lesson devised by Portuges and Feshbach (1972) is an individually administered task that includes both intentional responses (e.g., that the elephant is the largest animal in Africa, the cheetah is the fastest) and incidental ones (either gestural, such as arm-folding or pointing to one's head, or verbal directions such as "think carefully" or "listen hard"). However, the manipulation is presented on film, which insures stimulus constancy and makes it easier to administer; pilot study showed no differences in effects using live rather than filmed models. The child is told he will watch the film and then be asked to teach the geography lesson himself.

After watching the film, the child is presented with the lesson props (a background map, pictures of animals) and is asked to teach the geography lesson. Scoring includes the number of intentional and incidental responses occurring in the child's presentation.

- c. The content of the geography lesson is recommended as more relevant to the kind of school learning task children typically face, and it requires fewer and less elaborate props. An experimental situation of this sort is recommended, with a filmed stimulus presentation for purposes of stimulus constancy (interpersonal affect during the task might otherwise be the single most important determinant of learning). Further, with a filmed stimulus, children could observe it in small groups after which they could be taken to separate testing areas for scoring of learned responses.

The procedures for devising experimental responses should be adapted from the Postman Game. That is, the same number of intentional and incidental responses should be involved in the presentation. Further, the intentional ones should be explicitly intentional, e.g., the "teacher" should underscore the importance of remembering that the elephant is the largest animal in Africa.

In contrast, the incidental cues should be clearly incidental: if the child knows he is going to have to enact the teacher role in the lesson, he may think it relevant to remember the exhortative phrases used by the stimulus



teacher ("think carefully," "listen hard," etc.) in the film but the gestural cues used in the lesson and in the Postman Game are clearly incidental and are the sort that should be adopted. Scoring should be based on procedures used in the geography lesson--i.e. only one scoring situation is needed.

- d. Portuges comments<sup>1</sup> that although demographic variables of elementary school teachers are not subject to the child's control, they often influence what the child learns. Use of a set of filmed stimuli would allow crossing teacher and child ethnicity. In this way it could be determined whether teacher ethnicity influenced dependency and thus influenced learning of intentional and incidental responses. Such a procedure would be relevant for the present evaluation study because Head Start teachers are more often of the same ethnic background as the child than are public school teachers. The change in teacher variables from the Head Start to the public school situation could counteract the decrease in wariness found by Zigler and Butterfield (1968), which could be researched systematically using ethnicity of the filmed stimulus in a subsample study.
  - e. However the stimulus situation is arranged, the outcome variables are incidental responses and intentional responses learned; derived scores include total responses learned and proportion of intentional to total responses learned. Among these, proportion of intentional responses learned is regarded as the most important index of relevant cue selection and has been found by Ross (1966) to discriminate high dependency children. Both the Ross and the Portuges and Feshbach approaches have been used with preschoolers, and the Portuges and Feshbach study involves both Black and white children; in the latter study, white children were found to learn more of both kinds of behavior. It is expected that Head Start children would be more able than control children to select and learn relevant responses.
2. Reinforcement style, another learning-related outcome, is recommended for assessment using either a concept-switching or a discrimination task. The Zigler and de Labry (1962) concept-switching approach proposed for evaluating *boundary elasticity* in response to nonpersonal problem-solving, is described in the text of Chapter 5 and is not detailed here. The subsample study of reinforcement style could be incorporated into that paradigm. Should integration prove more complicating than

---

<sup>1</sup>Dr. Stephen Portuges is Research Director of Los Angeles Psychiatric Services, Los Angeles, California.

simplifying, a discrimination task should be used instead. Pilot work in either case should finalize procedures; specifically, a method of holding social reinforcement constant while giving tangible rather than informational reinforcement to treatment subjects and not control subjects (crossed-design) must be worked out. It is expected that the task can be group administered.

- a. The kinds of tasks used involve either discrimination learning or concept-switching (Terrel, Durkin, and Wesley, 1959; Terrel and Kennedy, 1957; Zigler and de Labry, 1968; Block and Block, 1973). Terrel's work makes use of three pairs of stimuli (cubes, cones, and cylinders), one large and one small of each, presented in random order and position. The subject task is to learn that the "correct" stimulus choice is the larger, the acquisition criterion being 9 correct choices out of 10. Block and Block use an acquisition-and-concept-switching task in which subjects first guess whether a red or green light will go on. After subjects have learned that the colors always alternate, the green light is switched on continuously until the child learns the new pattern. The score on the acquisition task is number of trials to criterion (8 consecutive correct guesses), concept-switching scored as the number of the last trial on which the child made a guess of red. The concept-switching task used by Zigler and de Labry (1968) is much more complex and is recommended only if concept-switching per se is of independent interest. Either of the other two tasks seems equally feasible.
- b. Reinforcers studied include a prize, candy, praise, a token that can later be "cashed in," reproof, and information that one has succeeded in doing the task correctly (Terrel, Durkin, and Wesley, 1959; Terrel and Kennedy, 1957; Zigler and de Labry, 1968; Block and Block, 1973). For the purposes of this study, prizes and candy can be regarded as equivalent material reinforcers. Use of tokens and negative reinforcers will not be considered, since they involve learning dimensions not related to the theoretical area of interest here. However, praise is a social reinforcer closely related to the notion that one is correct, and these two latter sorts of reinforcement must be compared with one another and with the efficacy of material reinforcers.
  1. In the Terrel approach correctness information is provided mechanically by a signal light. In the material reinforcer condition, the correctness signal flashes and the child also receives an automatically dispensed piece of candy. While this manipulation is uncontaminated by any social reinforcement, it is not very representative of the situation in which school-like tasks are usually performed.

- ii. In the Zigler and de Labry (1968) study, informational feedback is provided by a live examiner who seems interested in whether the child will be able to perform correctly (e.g., Block and Block, 1973). In Zigler's work, this reinforcement condition is contrasted with the condition in which a live examiner emphasizes that the child will receive a prize if he performs correctly. Thus, social reinforcement seems to accompany both kinds of reward. But in this situation the child receives one-to-one attention from the live examiner, a situation unlike most school learning tasks where the teacher gives feedback to a group of children at once.
3. Range of response repertoire given nonpersonal stimuli is suggested for measurement in a situation in which initial conditions do not determine a priori the nature and number of alternative responses. Thus an assessment of the nature and number of responses should provide an index of response resiliency. The situation is one in which many appropriate responses may be made and the stimulus conditions do not rule any of them out. It is, then, a nonfrustrating situation susceptible to multiple correct solutions. Pilot work is required to determine the extent to which this outcome class is representative of the construct of resiliency and to select the most feasible means of measuring it. The following measures are suggested for investigation and administration on a subsample basis:
    - a. The Sigel unstructured object-sorting task where subjects are asked to produce one or more sorts (described in Block and Block, 1973) might be appropriate for this purpose. No single (or, no two or three) a priori sorting principles are immediately apparent in the object assortment, so grouping must reflect the child's initiative in dealing with the materials.
    - b. The stimulus materials used by the Blocks to study parent teaching strategies (Block and Block, 1973) could also be used to study the generation of multiple correct solutions. Materials are either varied sizes of blocks or posts, and the goal of the task is to produce a large block or post matching in size a given criterion object. There are many ways of combining the stimulus materials to produce an object of the required dimensions; scoring

might reflect a number of different solutions generated in a given time period and degree of prodding required.

- c. Finally, the Block and Block "divergent thinking" and "unusual uses" tests seem to involve most of the features characterizing resiliency under nonfrustrating conditions. That is, these tests present familiar stimulus objects to children, ascertain that the ordinary use or nature of the object is known, and then ask the subject to think of other uses to which the object can be put or divergent ways of regarding it. While the measure focuses on response resiliency (given that a familiar response is ruled out), the extent to which it depends on verbal ability is problematic and needs investigation.
4. Attitudes toward self and school are the final outcome classes about which procedural details are provided to guide subsample studies. Measures for these attitudes are included in the major battery for the entire sample, but more confidence will reside in the results if they are substantiated by conclusions formed on the basis of intensive interviews. The attitude interviews recommended for subsample study, then, are intended to yield criterion variables against which results from other attitude measures can be weighed.
    - a. An individual interview should investigate the most important school attitudes. Pilot investigation is needed not only to decide on the best method of posing questions and the optimal number of questions usable, but also to decide what are the most important components of school attitude. Stearns (1971) suggests that it is important to ask whether the child enjoys school currently, although few researchers directly pose this question. Second, questions of success expectancy are relevant. Kagan (1971) points up the need for lower status children to want to do well and to believe that they can do well in school, summarized as a need to feel that intellectual skills are appropriate to their own identity. Finally, Kagan (1971) and Sarbin (1964) underscore the influence of beliefs about the "reinforcingness" of the social ecology: the child needs to believe that planning and effort-taking are worthwhile, that his progress (however small) will be recognized and rewarded by significant others (teachers and parents).

The attitudinal components mentioned above should not be regarded as exhausting the domain of critical school-related feelings, beliefs, and values. They do provide examples of such items. It would be expected that Head Start may well favorably affect such attitudes. The danger of the one-to-one verbal interview is that the lower status child's apprehensiveness toward the situation may mask real effects (Labov, 1970; Williams, 1970).

Suggestions below concern techniques the pilot study should explore for alleviating this problem.

1. Labov (1970) is perhaps the most successful interviewer of young respondents, and it is recommended that his methods be carefully explored. With young children, Labov recommends first that the interview be conducted at home if possible, in the child's room. Second, he thinks the power difference between interviewer and child that detracts from communication is partly overcome if the interviewer sits on the floor with the child and shares potato chips with him. Further, Labov has had considerable success getting young children to respond when he interviews two at once, which alleviates the children's shyness and permits them to outnumber the potentially threatening interviewer. Finally, it is seen as important to have the interviewer come from the same ethnic and status group as the child and, if possible, from the same neighborhood.
- ii. Other techniques might be borrowed from psycholinguistics, including two mentioned in Cazden's reviews (1966, 1967). One is the Deutsch telephone interview method, whose success seems to vary in relation to the topic of the interview. Another is letting the child select school photographs and then having him discuss the pictures. It has been found that when the child himself has selected the picture stimulus, he is much more willing to discuss it. The choices themselves might provide an unobtrusive attitude index.
- iii. Social psychologists (Jones and Gerard, 1967) have had some success interviewing young children on their attitudes toward comics by pretending the purpose of the interview was to make recommendations regarding still younger children (e.g., would it be okay for your little brother to read \_\_\_\_\_ comics?). A similar approach could be tried by asking children in their first year of public school what the incoming population of students should be told about school (e.g., whether it was fun, whether they should wait a year and go when they were a little older, etc.). It was recommended by a nine-year-old<sup>1</sup> that third-graders in the same school system might be a fruitful source of questions--children at this age are sufficiently verbal and can express themselves fairly well; but they are close enough to their first school experiences to remember what was salient, how it influenced them, and so on. In addition, they are familiar with the colloquialisms of the neighborhood and can provide interviewers with useful phrasings.

---

<sup>1</sup>This suggestion comes from Kara L. Bikson (the writer's daughter).



- b. Exactly the same procedural interview techniques as are recommended for the school attitude interview apply in relation to the self-attitude interview; in fact, it is assumed that the same individual interview session will investigate both of these attitudinal outcomes. Pilot investigation will here too have to determine not only the best method of putting questions to children, but also what are the most important self-attitude contents to evaluate. The discussion below focuses on sources on which the research staff might rely in generating interview topics.
- i. Two suggestions for interview instruments come from recent research by Powell.<sup>1</sup> Dr. Powell has successfully used the Piers-Harris Children's Self Concept Scale with children as young as third grade (Piers and Harris, 1964). At that grade level the verbally administered self-attitude scale is internally reliable (coefficients in the 90's) and has test-retest stability (coefficients in the 70's after four months); evidence of concurrent validity is also impressive (Piers and Harris, 1969). Powell thinks that the Piers-Harris scale could be revised downward so the item content would be applicable for kindergarteners and first-graders. It is recommended that this suggestion be implemented in pilot work designed to determine whether the modified version retained reliability and validity for Head Start and post-Head Start children.
  - ii. Of greater value, according to Powell, would be a revised version of the Tennessee Self Concept Scale (Fitts, 1964). This self-description inventory has been found reliable and valid among high school students, but the item content is almost never applicable for children in the age range of the prospective research population. What is most impressive about the Tennessee scale and worth trying to replicate for younger populations is that it distinguishes classes of self-concept, assessing self-evaluations relative to the social self, academic self, family self, moral self, and physical self. Clearly these are distinguishable dimensions of the self-construct, and Powell's use of this scale with older students indicates that, at least with respect to several of the dimensions mentioned, minority subjects express self-evaluations equalling or surpassing those expressed by middle status white subjects. Clearly it would be desirable

<sup>1</sup>Dr. Gloria Powell is a child psychiatrist holding positions at the Martin Luther King Hospital and Charles Drew Post-Graduate Medical School (Los Angeles) as well as at the University of California at Los Angeles Neuropsychiatric Institute.

to differentiate these dimensions of self-concept in an interview situation, in order to determine the aspects most influenced by Head Start experience.

111. A final instrument suggestion is the set of Self Observation Scales (SOS) just developed and published by National Testing Service. The 40-item instrument for early primary grades has subscales indicating self-acceptance along with four other important attitudes. The scales have not been investigated for use specifically with low-SES and minority populations, however, and while they appear to be reliable no evidence of external validity of any sort is presented in the technical reports. The interview-research staff should examine the scales with these questions in mind.



Appendix G

LITERATURE SURVEY ON INDEPENDENT VARIABLES

CURRICULUM VARIABLES

- o Strong emphases on language (System Development Corporation, 1972b)
  - Strong emphasis on Language Program is negatively related to Stanford-Binet scores of high IQ children.
  - Strong emphasis on Language Program is negatively related to PSI and Birch-Verbal response scores of young children.
  - Strong emphasis on Language Program is positively related to PSI and Birch-Verbal response scores of older children.
  - There is an overall positive effect of Birch Spontaneous Scores for all children (may be indicative of verbal fluency).
  - In her review of preschool impact, Stearns (1971) found that use of language was positively related to increases in ITPA scores.
  - An evaluation of Dawe's preschool program emphasizing language development has found that access to books, hearing poems and stories, going on trips and excursions, discussing pictures, and training in words and concepts have produced significant gains on the Stanford-Binet and Van Wagner Reading Readiness Tests. Also affected are the length and complexity of sentences uttered (Dawe, 1942).
  - IQ gains have been reported from programs that focus on the area of a child's weakness, whether motor, language, or visual (Coffman and Dunlop, cited in A. Butler, 1970).
- o Emphasis on socialization skills (SDC; 1972b)
  - Emphasis on socialization has its strongest relationship in the affective/social domain (enhances the child's adjustment to Stanford-Binet test conditions and Birch verbal responses); also, it is positively related to PSI.
- o Emphasis on child independence and self-care (SDC, 1972b)

- Independence and self-care are positively related to several measures of cognitive and affective behavior. However, the relationship becomes complex. There is no linear relationship between this variable and the Binet or PSI for high IQ children. A relationship is almost totally absent from mid- and low-IQ individuals.
- In a study of Follow Through classrooms, up to a moderate level of pupil freedom is positively related to student growth in complex-abstract thinking; beyond a certain point, increased freedom leads to less growth. In contrast, simple, concrete growth shows no relation to pupil freedom (Soar, 1971).
- o Personal and social development effects (Dittman et al., 1970-71; Emmerich, 1971)
  - These effects are positively related to specific personal and social behavior in the earlier period of the year but tend to level off or become less regular through mid-winter and spring (may have implications for testing schedule).
- o Emphasis on structured vs. unstructured curricula (Stearns, 1971)
  - Unstructured programs are positively related to poor IQ scores.
  - Unit-based instruction vs. cognitively based curriculum vs. language-training curriculum show no differential relationship on cognitive measures.
  - Re-analysis of earlier studies by Smith and Bissell (1970) indicates that, overall, highly structured programs are more effective in producing cognitive benefits than less structured. However, when these highly structured programs were analyzed in relation to SES levels, the more advantaged of the lower class children gained as much or more from *unstructured* programs, whereas the less advantaged of the lower class gained more from structured programs. These findings supplement prior preschool data from HSPV.

- "Less directive models" have a positive relationship to Stanford-Binet and PSI scores of children in the second year of Head Start. "More directive models" have positive relationship for first-year children on Stanford-Binet and PSI. HSPV data do not show consistent results by ethnicity and SES across tests or years. Therefore, these data would not help predict which children will be helped most by what model.
- o HSPV models that are "more academic" (M. Smith, 1973b, Huron Second year report),
  - These models are more effective in transmitting academic skills, as measured by PSI, WRTR (letters), WRTD (numbers) and ETS.
  - In comparisons of the HSPV models, non-Planned Variations, and control groups, both HSPV and NPV groups had positive gains on all tests (PSI, WRTR, WRTD, ETS), except PPVT. The control group had smaller gains by comparison.
- o Interaction between curriculum and measures (L. Miller et al., 1971)
  - In comparisons of children in the Montessori, Cognitive Discovery, and Prescriptive models, there were modest gains relative to the traditional model (no treatment controls). The question becomes a longitudinal one to see whether different curricula approach the same final goal at *different rates* or at the same rates.
- o Effects of structured transitions between preschool and kindergarten (L. Miller et al., 1971)
  - The following results are reported at the end of the *second year* of a three-year comparison of four pre-kindergarten programs (Bereiter-Englemann, DARCEE, Montessori, and Traditional). These program styles were used with 4-year olds in Head Start. Kindergarten experience was varied; some children entered Follow Through Kindergarten, the remainder entered Regular Kindergarten, a non-academic program. The two kindergarten programs *did* differ significantly in many of the ways predicted:

- \* Overall, Follow Through effects on outcome variables, with initial level controlled for, are fewer and of smaller magnitude than hoped.
- \* The difference between the means for Follow Through and Regular Kindergarten on the PSI is a significant one. However, since it is a difference of only about 2 points, one might question the psychological significance of this small amount.
- \* It could be tentatively concluded that for children without Head Start or those who have not had a Head Start program emphasizing *language training*, Follow Through Kindergarten would be an advantage in respect to language ability.
- \* On one measure, *arithmetic*, Follow Through did substantially amplify the differential gains obtained in Head Start. This result suggests that the benefits of Follow Through are more likely to be found in such achievement measures as school readiness and school achievement tests.
- \* There was a decrease in *persistence* for DARCEE and Montessori children who experienced Follow Through Kindergarten, since Follow Through teachers were clearly reinforcing persistence. It was hypothesized that this may be an example of faster extinction of learned responses after continuous reinforcement since in the test situation children receive no reinforcement. The supervisor of Follow Through commented that the teachers had observed their children in other test situations "waiting for reinforcement before proceeding."
- \* The highest scores on *Embedded Figure* were obtained by children who had DARCEE Head Start followed by Follow Through Kindergarten. This is especially interesting because there were no program differences on this variable at the end of Head Start, suggesting that there may have been Head Start program effects that were not measured, despite the fairly large battery used.
- \* Stable Head Start program effects, regardless of type of kindergarten, may represent modifications at the

four-year-old level that continue to influence behavior despite wide variations in subsequent educational experience.

- \* Regardless of the type of kindergarten, children from Bereiter-Englemann Head Start were still manifesting a decided tendency to resist distraction at a task. The controls had lower scores. Traditional children remained low in curiosity. However, both Bereiter-Englemann and Traditional Head Start children remained below the level of control children who had not had any Head Start in devising alternative solutions to a problem, as shown by scores on Inventiveness. Montessori and DARCEE children remained high in Inventiveness, regardless of kindergarten. This measure appears similar to Guilford's definition of divergent thinking. One might speculate that Bereiter-Englemann children, having been drilled to give the correct answer, thereby lost a certain flexibility in providing alternative solutions; however, if this explanation is correct, there must be some *other* reason why children from Traditional Head Start were also very poor on divergent thinking.
- \* For preschool and kindergarten children, it may be desirable to provide somewhat different programs or program components for boys and girls. In view of the main effect of sex on the Binet, sex differences could be due to differences in intellectual maturity. Temperamental or experiential differences may account for different reactions of the two sexes. From a report based on monitoring videotapes and observing classes, females, in general, were more attentive at this age. More teacher attention was directed toward females in most classes in the Head Start year. However, males may have been participating less.

#### PROCESS VARIABLES

##### Teacher Background Variables

- o Teacher's paid experience with disadvantaged youth (SDC, 1972b; Stanford Research Institute, 1973)

- This variable was negatively related to PSI scores of primarily non-southern children who have low to middle IQ.
- There was no relationship with PSI scores of southern children.
- This variable was negatively related to attitudes of parents of high-IQ children, primarily in an urban setting.
- o Level of teacher's general educational preparation interacts with demographic variables (SDC, 1972b)
  - There is no overall effect on the Stanford-Binet scores of Head Start children.
  - Higher educational levels of teachers in mostly urban settings are negatively related to PSI scores, children's ability to adapt to test conditions, and observed work responses and verbal responses on Stanford-Binet.
  - The level of a teacher's education has no effect on parent attitudes nor on scores for cognitive measures among non-urban children.
- o Teacher qualifications
  - In a study of grade school teachers, Shim (cited in A. Butler, 1970) found that experienced noncertified teachers with less than a 2.5 GPA were more successful in raising the IQ and language achievement of students of below average intelligence than less experienced certified teachers with better grades. However, there was no relation of the above teacher characteristics to achievement in students of above average intelligence.
  - In grade school, Yee (1968) found a negative correlation between years of teacher experience and lower class students' attitudes toward the teacher. On the other hand, the sex of the teacher (male) was positively related to positive attitudes toward the teacher in both middle and lower class students.

### Classroom Atmosphere Variables

The classroom atmosphere variables include Teacher Behavior, Child Behavior, and Class Activities and/or Settings. Each subsection includes a survey of the literature and a summary of the studies.

- o Teachers who gave unconditional warmth and support
  - This characteristic was positively related to children's adjustment to the school setting (higher SES sample used).
- o Teachers who were highly vs. moderately encouraging
  - This characteristic was negatively related to IQ growth (Eisenberg, cited in Grotberg, 1969b).
- o Motivational considerations
  - When motivation increases through a decrease in wariness of adults, IQ gains are noted (Zigler and Butterfield, 1968).
  - Social reinforcement from important adults is positively related to motivation in intellectual tasks and the development of autonomy (Zigler, 1970).
- o Teachers who placed high value on intellectual activity
  - This characteristic was positively related to intellectual growth.
- o Teachers who placed high value on property rights and care of materials
  - This characteristic was negatively related to IQ growth.
- o Teachers who placed *heavy vs. moderate* emphasis on self-confidence and self-concept
  - This characteristic was negatively related to intellectual growth (Eisenberg, cited in Grotberg, 1969b)
- o Teachers who are abstract and complex (less structured, less punitive, more resourceful and flexible) (Grotberg, 1969b)

These characteristics were:

  - Positively related to increase in child's self-esteem.
  - Positively related to child's involvement in activities.
  - Positively related to increases in child's achievement.
  - Negatively related to child's concrete responses.



- Positively related to child's cooperation.

Note: Only 8% of teachers were found to be abstract.

- o Providing teacher feedback, as opposed to no feedback, about children's performance and about interactions in their classroom
  - No overall effect in the children's performance found.
- o Frequent teacher-child interactions (school data)
  - This characteristic was positively related to adjustment to school.
- o Teacher's instructional pattern
  - Amount of exposure to situations of adult question, child response, and adult feedback is negatively related to kindergarten and first grade achievement and positively related to absenteeism (Stallings et al. (in press), cited in Brandt, 1972).
- o Teacher activity level (amount of verbalization or number of communicative episodes (Moore, 1971; Smothergill et al., 1972, cited in Cazden, 1972)).
  - There were two tutorial treatments: "patterning" (teachers elicited from children particular language forms) and "extension" (the teacher responded to children's comments by modeling such elaborated use herself).
  - There were no treatment differences on the sentence imitation tests; for the other measures (WPPSI and measures of communication effectiveness) the patterning program was superior. In Smothergill's research, results indicated that the elaborately taught group gave more task-relevant elaborations and performed better from pre- to posttest on the verbal similarities task and on the story telling task. The non-elaborative group gave more spontaneous directives, many of which were attempts to get the teacher's help and attention. The groups did not differ significantly on three non-verbal problem-solving tasks or on time spent on teaching activities.

In summary, of the two variables of teacher elaboration and teacher elicitation of child elaboration, teacher elicitation is specifically responsible for greater elaborate behavior by the child.

- o Teacher's quality of cognitive input (SDC, 1972b)
  - This behavior does not seem to have an overall effect on PSI or Stanford-Binet. However, with a high IQ group of children, this behavior is positively related to the parent's feeling of alienation.
- o Teacher's use of physical control (SDC, 1972b)
  - This behavior was negatively related to Stanford-Binet and most of the subset scores of PSI. The highest scores were received when no physical control was used.
- o Teacher criticism (Soar, 1973)
  - This behavior was negatively related to pupil growth in reading, creativity, and vocabulary.
- o Teacher's instructional style (Lamb, Ziller, Maloney, cited in A. Butler, 1970)
  - If the style was more *abstract*, the children gained more in self-esteem, identified closely with their mothers, and perceived themselves as similar to others. If the style was more concrete, the reverse was true.
  - The extent to which the teacher (1) requires 100% responses from all students, (2) corrects errors by repeating the entire task and retesting the child, and (3) follows a specified lesson format is positively related to student cognitive gains (Rosenshine and Furst, 1973).
  - Highly guiding child's activities was found to be positively related to student constructive behavior in the face of failure, more participation and leadership, and less destructive behavior (Thompson, cited in A. Butler, 1970).
- o Teacher's exercise of power (Prescott, 1971)
  - High exercise of power is positively related to the number of lessons that are taught to the children, to the use of individualized teaching, and to whether the teacher is highly encouraging or highly restrictive.

- Low exercise of power is positively related to the teacher's large amount of non-communicative behavior, to greater attention to children's physical care, to the use of *routine* encouragement or restriction, and to group teaching.
- o Teacher style (as measured by Observers Rating Form along such dimensions as depending, irritability) (Linn, 1966)
  - Teacher style was positively related to child's PPVT.
  - Teacher style was positively related to Scale II of Pre-school Inventory (personal and social responsiveness).
- o Teacher indirectness as measured by Flander's Interaction Analysis (asking questions, accepting pupil's ideas, praising, encouraging) (Soar, 1973)

This behavior was:

  - Positively related to subject matter achievement (3rd to 6th grades).
  - Positively related to student attitudes toward school and teacher.
  - Positively related to student growth on abstract learning task.
  - Negatively related to growth on concrete learning task.
  - Positively related to growth over summer vacation.
- o In their review of teacher variables, Rosenshine and Furst (1973) found that teacher's clarity, flexibility, enthusiasm, and task orientation are consistently positively related to student achievement.
- o Small group instruction (SRI Follow Through Observational Data, 1973)
  - There is a significant relationship between high test scores and small group instruction and a stimulus-response-feedback interaction (correlations are all above .43).
- o Peer group effects (Coleman (1966) school data)
  - There was a strong relation between pupil's achievement and the educational backgrounds and aspirations of other students in the school. In other words, the composition of the student body was found to be more important in

predicting pupil achievement than such characteristics as school facilities, teachers, or curricula. This effect may be interesting to study in Head Start children. Even though several studies refer to lower class preschool children's inability to relate to the breakdown of authority, no study has been done investigating the Head Start children's perceptions of an authority figure and how this may interact with peer effects. Although many authorities will agree the teacher is the most important determinant of structure in the preschool classroom, it has *not* been illustrated that this is the case for all children from different backgrounds.

- Smith (1968) found that preschool children produced more complex and lengthy sentences when in a peer situation than in a teacher-child situation.
- Radin and Weikart (1967) found that participation of other children in the home-based complement to his nursery program was negatively related to IQ gains on the Stanford-Binet.
- Soar (1971) reports that expression of negative feelings among students relates negatively to subject matter achievement.

o Attendance

- Studies investigating attendance (greater number of days correlated to scores on certain measures) have been largely unsuccessful. Replication of the studies, using larger sample sizes, have been deemed essential.

Note: The attrition rate of children has been shown to be an important outcome variable when experimental groups of children are compared with children used as controls. In a study cited by Erickson (1969) attrition among control families approached 50%, while among Bereiter-Englemann preschoolers and traditional preschool families it approached about 17%. Unfortunately, no data are given on the IQ and achievement characteristics of this attrition, but there is a possibility that it was not random.

- o Child behavior including independence, task persistence, and cooperation (SRI Follow Through Observational Data, 1973) (*independence* is defined as children engaged in a *task* without an adult)

- In classrooms where teachers allow children to select their own seating and groups part of the time, where a wide variety of activities are available, and where there is an assortment of audiovisual and exploratory materials available, children were more independent.
- Fewer independent children are observed in classrooms where textbooks and workbooks are used frequently. Adults who ask more direct questions regarding the subject matter are also less likely to have independent children.
- In classrooms where adults praise children a lot (the variable describes praise in general, not for specific tasks or achievement), children are less likely to be independent. This negative relationship is a very high  $-.60$ .

*Task persistence* is defined as children or a child engaged in self-instruction over a designated period of time.

- The highest positive relationships indicate that task persistence occurs most often when textbooks and workbooks are used in the classroom.
- Where adults instruct one child at a time, the children are also likely to be more task persistent.

*Cooperation* is defined as two or more children working together on a joint task.

- This kind of cooperation is more likely to be found in situations where a wide variety of activities occur throughout the day, when exploratory materials are available, and where children can choose their own group. If the adults interact with two children asking questions and making comments about the task, the children seem to be encouraged to join each other in cooperative tasks. When textbooks and workbooks are used a great deal, the children are not likely to cooperate. (There is a strong negative correlation of  $-.52$ .)

- o Child's involvement in the classroom (has also been linked with attentiveness)
  - This behavior was consistently correlated with a child's achievement on the tasks attended to.
  - This behavior was highest in classrooms observed in the more pupil-controlled or pupil-centered settings, especially those that entailed cooperative interaction among the children, with or without the teacher as a participant in this cooperation.
- o Effects of testing periods and length of time in program
  - Up to 15 days (SRI-PV 1971)
  - 15 to 30 days
  - 30 days or over
  - Regardless of previous Head Start experience, there was no overall systematic increase in initial scores over all children due to the effects of time elapsed before initial testing. Significant differences occurred between children with previous Head Start experience who were tested in less than 15 days and those tested more than 15 days but less than 30 days after the start of classes on both preacademic and general cognitive measures. The group tested more than 30 days after start of classes had lower mean scores than the second group. Academically oriented sponsors have claimed that delays in testing prejudice evaluation by raising "initial scores" since a significant amount of learning is achieved by the children in the first six weeks of the program.

Note: An interesting artifact of Head Start programs, in view of the previous paragraph, was the answer to the question "when are centers expected to complete enrollment, even if there was a 'best time' to begin testing?" A small survey of the Head Start data indicated how various programs interpreted "the completion of enrollment." Some programs indicated completion by:

end of August

end of September

early October

end of December

June 1 (assumed to be the following year)

mid-November

next spring

- o Effects of structured transitions within classroom daily activities (Prescott, 1971)
  - Comparing centers with home-based care indicated that there is a higher percentage of health-related activities (such as sleeping, eating, toileting) in the center (24%) than in home-based care (3%).
- o Low vs. high level cognitive activity in a Follow Through evaluation
  - Soar (1971) found that activities of a lower cognitive level (e.g., rote memory, identification) are positively related to pupil growth in abstract thinking while activities of a higher cognitive level (e.g., application, analysis, evaluation) are negatively related to such growth. Involvement of pupils in greater amounts of complex thinking skills appears not to be functional.
- o Class activities log sheet (Brandt, 1972)
  - The time is recorded to the nearest five minutes by the teacher or an aide when a shift occurs from one activity and categorical estimates are checked by the model types of (1) motor activity, (2) grouping pattern, (3) activity selector, and content emphasis which prevailed during the previous activity. When daily activities were logged in ten summer Follow Through classrooms, teachers were found to make or help make over three-quarters of the selections of children's activities, and content had an intellectual emphasis almost half the time. As compared with studies in day care, amount of free choice was highest in home-based situation.



o Follow Through Classroom Process Measurements (Soar, 1971)

(Sample: seven programs, at least 8 classrooms)

- Results present a portion of the evaluation of the planned variation of Project Follow Through.

(Program Discrimination)

- \* Highly significant differences in classroom behavior are associated with differences in sponsorship.
- \* The instrument, *Teacher Practices Observation Record*, discriminates programs significantly for five of the six factors measured.
- \* There appear to be differences in the degree to which sponsors have been successful in implementing their objectives, and it seems probable that stresses in communities that are reflected in the school have negated the sponsors' efforts in some cases.
- \* It also seems likely that programs differ in the difficulty of implementation. Nevertheless, the success of sponsors in producing classrooms that reflect their objectives as they are measured in this report seems striking (pupil growth data).
- \* The total test battery can be broken down into largely independent subscores representing rather different kinds of learning. What has been called *simple-concrete* learning seems to require little but memory; *skill* represents the acquisition of the traditional academic skills; *complex-abstract learning* seems to require complex information processing, solving complex problems, or comparing complex figures. These different classes of scores do not relate strongly with others; further, they seem to respond differently to the dimensions of classroom behavior and to different programs.

It seems especially important to find some fairly independent classes of measures (since the total battery used in the evaluation is more heavily weighted with items that represent the *skill* measure) so that a sum

for the total battery would favor programs that emphasize teaching skills. When these measures are used to examine a group of pupil subgroups differing in ethnic group and socioeconomic status, the minor differences in growth in either *concrete* or *abstract* measures are so inconsistent, it would lead one to conclude that there are no real differences. Subgroups start at different levels and finish at different levels, but their scores grow at similar rates for both abstract and concrete measures. These results appear to agree with other authors who have found that the major differences between social status groups in the amount of academic growth that took place during elementary school occurred during the summers, rather than during the school year.

- o Relations between observational data and pupil growth (Soar, 1973)
  - In contrast to the use of observational data in program discrimination, the relations of the observational measures to measures of pupil growth are scattered and often inconsistent. Several problems contribute to difficulty in drawing dependable conclusions.
    - \* Growth is the problem being studied, but growth measures are much less reliable and correlate much less strongly with each other and with other measures than do scores that represent standing at a point in time.
    - \* Thorndike (1966) estimates that the correlation of a child's standing at the beginning of the year with growth during the year is probably no more than + .10. In contrast, correlations between standings for elementary pupils commonly approach + .70.
  - Variables reflecting negative emotional climate fail to relate to measures of pupil growth. In contrast, two factors reflecting positive involvement of the teacher with pupils relate to growth.

- When the measures as a group and their relations with pupil growth are inspected, factors that discriminate programs most strongly and that represent major portions of the variance in the observational data are *not* the ones that relate strongly to pupil growth. (There was a tendency for the above relations to be non-linear; also, the number of classrooms within any grade levels group was too small to permit reasonable test of the hypothesis.)
- Observation methods appear to be significant discriminators of at least some program objectives, even though the relations between observational data and pupil growth are less clear. (Measures: Complex-abstract, skill, and simple-concrete subgroups of pupil growth measures were created in this study.)
- o Observer Effects (Samph, cited in Soar, 1973)  
(Data on 4th and 5th grade classrooms)
  - Five variables from Flanders Interaction Analysis were tested for significance of change (all comparisons were in terms of deviations of each teacher from her own ideal). A comparison was made of base-line data collected when a previously scheduled observer was present in the classroom. Significant change was found for two of the five variables: the amount of praise produced by the teacher increased when an observer was present, and the amount of criticism decreased. In each case the difference between means for the control and experimental conditions was about three-quarters of a standard deviation. This is the variability of differences between observed and ideal behavior for individual teachers, and it probably is much smaller than the variability of behavior across teachers. None of the other three variables showed significant change (p. 33).
  - Overall, it seems reasonable to assume that teacher behavior does not change greatly as a consequence of the presence of an observer during classroom activity.

### Organizational Variables

Organizational variables include equipment and materials, physical facilities, sponsorship, and parent participation.

- o Cognitive learning materials (SDC, 1972)
  - These materials are positively related to *Birch Work* response scores and scores on Stanford-Binet. They were negatively related to scores on *Animal House*, which is supposed to have higher cognitive values.
- o A multicultural primer with natural, familiar speech patterns produced high interest, increased verbal competence, and increased word recognition in school-age children; Black children benefited the most (Whipple, cited in Linn, 1966).
- o "Large muscle" equipment (SDC, 1973)
  - This equipment was positively correlated to higher scores on Stanford-Binet and PSI.
- o Size of center
  - At least in many Day Care Studies (Prescott, 1971; Handler, 1970), the size of the center is positively related to the program quality. For instance, in large centers (serving 60 children or more) there is more emphasis on rules and routine. Teacher control and restraint was 2-1/2 times greater and teachers' behavior was more neutral and distant. In small centers, the results were the opposite.
  - It is generally accepted that small centers show greater program variations and higher turnover rate in professional staff.
  - Turnover rate (impact of mobility) of staff has little effect on quality of Head Start programs in general. There is more effect on individual components of Head Start programs--i.e., social services and Health (Booz-Allen, 1972).
- o Differences in full-day vs. part-day programs
  - Part-day (3 to 6 hours) programs increase parent participation.

- Full-day programs (more than 6 hours) reveal the difficulty in scheduling times for staff members to talk together. From available figures about 30% of the programs are operating on a full-day schedule.
- No studies have evaluated the effect of sponsorship on classroom process.
- o Parent participation (Stearns, 1971; SDC, 1972a)
  - Part-day programs seem to increase parent participation in the classroom.
  - Overall, parental involvement shows conflicting results as to increase in the child's performance. Unless that participation was fairly intense, where more responsibility was bestowed on the parent in the education of the child, few significant relationships were revealed.
  - Home-based programs improved parental attitudes toward "gains in self-confidence." Home-based training seemed to improve the durability of preschool gains.
- o Mother-child relationships as measured by Hess-Shipman Eight Block Sort (considered to be the microcosm of Head Start and PV processes)
  - These interactions were positively related to positive change in affective relationships and to the child's response to an abstract conceptual task.
- o Effects on younger siblings of the preschool age child

If mothers were involved in the Head Start program, there was an increase in IQ scores of the younger child (Stearns, 1971). A "diffusion effect" seems to be present, affecting the younger sibling where the intervention is directed at the older sibling. These data are preliminary and more studies need to be done. Studies of the overall effect of Head Start programs on family situations seem to reveal minor to zero effects (Stearns, 1971).

BIBLIOGRAPHY

- Bereiter, C., and S. Englemann, *Teaching Disadvantaged Children in the Preschool*, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1966.
- Booz Allen & Hamilton Management Consultants, *Retrospective Study of Employee Mobility in Head Start Programs*, Washington, D.C. Prepared for Office of Child Development, U.S. Department of Health, Education and Welfare, May 1973.
- Brandt, R. M., *Observational Methodology for Evaluation of Early Childhood Programs*, Journal of Research and Development in Education, Vol. 6, No. 3, spring 1973, pp. 94-108.
- Butler, A. L., *Current Research in Early Childhood Education: A Compilation and Analysis for Program Planners*, Washington, D.C.: American Association of Elementary-Kindergarten-Nursery Educators, 1970.
- Cazden, C. B., "Some Questions for Research in Early Childhood Education," in J. C. Stanley (ed.), *Preschool Programs for the Disadvantaged: Five Experimental Approaches to Early Childhood Education*, The Johns Hopkins Press, Baltimore, 1972.
- Cole, M. and J. S. Bruner, "Preliminaries to a Theory of Cultural Differences," in Ira Gordon (ed.), *Early Childhood Education: The Seventy-first Yearbook of the National Society for the Study of Education*, University of Chicago Press, Chicago, Illinois, 1972, pp. 175-176.
- Coleman, J. S., O. E. Campbell, C. T. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld, and R. L. York, *Equality of Educational Opportunity*, Washington, D.C.: U.S. Department of Health, Education and Welfare, 1966.
- Dawe, H. C., *A Study of the Effect of an Educational Program Upon Language Development and Related Mental Functions in Young Children*, Journal of Experimental Education, Vol. 11, No. 2, pp. 200-209, December 1942.
- Dittman, L. E., H. Kyle, *Head Start Planned Variations case studies*, University of Maryland, 1970-71.
- Emmerich, W., *Disadvantaged Children and Their First School Experiences: Structure and Development of Personal-Social Behaviors in Preschool Settings*, Princeton, New Jersey: Educational Testing Service, PR71-20. Prepared under Grant H-8256, Department of Health, Education and Welfare, 1971.
- Erickson, E. L., J. McMillan, J. Bennell, L. Hoffman, and O. D. Callahan, *Experiments in Head Start and Early Education: Curriculum Structures and Teacher Attitudes*. Washington, D.C.: Office of Economic Opportunity. Project Head Start, 1969.

- Grannis, J. C., "Autonomy in Learning: An Exploration of Pupil's and Teacher's role in Different Classroom Environments to Develop Criteria and Procedures for Evaluation in Project Follow-Through," excerpt from final progress Report, Teachers College, Columbia University, New York, February 1973.
- Greenberg, P., *The Devil Has Slippery Shoes: A Biased Biography of the Child Development Group of Mississippi*, New York: MacMillan, 1969.
- Grotberg, E. H., *Review of Research, 1965 to 1969*. ERIC No. ED-028-308. Project Head Start, Office of Economic Opportunity, Pamphlet 6108-13. Washington, D.C., 1969.
- Handler, E. O., *Preschools and Their Graduates*, Doctoral Dissertation, University of Illinois, 1970.
- Jackson, P. W., *Life in Classrooms*, New York: Holt, Rinehart and Winston, 1968.
- Linn, E. L., *The Socially Disadvantaged Child: Teacher Correlates*. Ph.D. dissertation, University of Texas, Austin, 1966.
- Miller, L. B., et al., *Experimental Variation of Head Start Curricula: A Comparison of Current Approaches*. Progress Report No. 9, Louisville, Kentucky: University of Louisville, March-May, 1971.
- Palmer, F. H., C. B. Cazden, and J. Glick, *Evaluation of Day Care Centers: Summative and Formative*. In E. H. Grotbert (ed.), *Day Care: Resources for Decisions*. U.S. Office of Economic Opportunity, Office of Planning, Research and Evaluation, 1971, pp. 442-57.
- Prescott, E., et al., *Assessment of Child-Rearing Environment: An Ecological Approach* (series), The Day Care Environmental Inventory. Pasadena, California, Pacific Oaks College, Prepared for the Children's Bureau, Office of Child Development, U.S. Department of Health, Education, and Welfare, R-219-(C6). November 1972.
- Prescott, E., et al., *Assessment of Child-Rearing Environments: An Ecological Approach*, Pasadena, California, Pacific Oaks College, Prepared for Children's Bureau, Office of Child Development, U.S. Department of Health, Education and Welfare, June 1971.
- Radin, N. and D. P. Weikart, *A Home Teaching Program for Disadvantaged Preschool Children*, Journal of Special Education, Vol. 1, No. 2, pp. 183-195, 1967.
- Risley, T., and M. F. Catalao, "Evaluation of Planned Activities: The Pla-Check Measures of Classroom Participation," in Davidson, Clark, and Hammerlynck (eds.), *Evaluation of Social Programs in Community, Residential and School Settings*, Research Press, (campaign (in press)).
- Rosenshine, B., and N. Furst, "The Use of Direct Observation to Study Teaching," in M. W. Robert, Travers (ed.), *Second Handbook of Research on Teaching*, Rand McNally, Chicago, 1973, pp. 122-183.



Soar, R. S., *Follow-Through Classroom Process Measurement and Pupil Growth (1970-71): Final Report*, Gainesville, Florida, University of Florida, June 1973.

U.S. Department of Health, Education and Welfare, Office of Child Development, Research and Evaluation Division, *Project Head Start 1969-1970: A Descriptive Report of Programs and Participants*, HEW-391, July 1972.

Weikart, D. P., *Preschool Programs: Preliminary Findings*, Journal of Special Education, Vol. 1, No. 2, pp. 163-181, 1967.

Weisberg, H. I., *Short Term Cognitive Effects of Head Start Programs: A Report on the Third Year of Planned Variation - 1971-72*, Cambridge, Massachusetts, Huron Institute, August 1973. (Prepared for Grant #H1926 from Office of Child Development, Department of Health, Education, and Welfare, U.S. Government).

Wood, S. A., *A Multidimension Model for the Observation, Analysis and Assessment of Classroom Behavior*, Journal of Research and Development in Education, Fall, 1970, Vol. 4, No. 1, pp. 84-102.

Yee, A. H., *Interpersonal Attitudes of Teachers and Advantaged and Disadvantaged Pupils*, Journal of Human Resources, Vol. 3, No. 3, pp. 327-345, 1968.

Zigler, E., *The Environmental Mystique: Training the Intellect Versus Development of the Child*, Childhood Education, Vol. 46, No. 8, pp. 402-412, May 1970.

Zigler, E., and E. C. Butterfield, *Motivational Aspects of Change in IQ Test Performance in Culturally Deprived Nursery School Children*, Child Development, Vol. 39, pp. 1-14, 1968.

Appendix H  
CATEGORIZATION OF COUNTIES ACCORDING TO  
METROPOLITAN/SPARSENESS DIMENSION

This appendix contains two lists:

- o A list of counties categorized by metropolitan/non-metropolitan distinctions, and
- o A list of central city counties.

The first list has ten codes. Codes 0-3 are all metropolitan counties, distinguished by dimensions of no relevance to the Head Start design stratifications. Since none of the metropolitan codes distinguishes central city counties, the second list should be used for categorizing Head Start centers as central city--not central city centers. The center is categorized as a central city center if it occurs in a central city county. Metropolitan, non-central city counties are those that occur in categories 0-3 on the first list and do not occur on the list of central city counties. Codes 4-9 are the non-metropolitan counties.

The first list, which is compiled by the U.S. Department of Agriculture, is continually updated. Prior to categorizing the Head Start centers along metropolitan/non-metropolitan lines, the contractor should obtain an updated list. David L. Brown, Population Studies Group, Economic Research Service, in the U.S. Department of Agriculture can provide the list.

The second list is derived from census data published by the U.S. Department of Commerce, *General Demographic Trends for Metropolitan Areas, 1960-1970*, Final report, Publication No. PHC(2)-1, U.S. Government Printing Office, Washington, D.C., October 1971, pp. 105-114.





CODE 2: 3003

00186



CONF 3: LESSON #100

[illegible]



HUMAN RESOURCES PROFILE COUNTIES

CODE 02 NONMETROPOLITAN URBANIZED ADJACENT

ALA	CALHOUN	IND	WAYNE	N.M.	SANTA FE	PA	CENTRE
ALA	DALLAS	IOVA	CLINTON	N.Y.	CATTARAUGUS	PA	CLEARFIELD
ALA	LEE	IOVA	JOHNSON	N.Y.	CAYUGA	PA	COLUMBIA
ALA	MORGAN	IOVA	MUSCATINE	N.Y.	CHAUTAUQUE	PA	CRAWFORD
ALA	TALLADEGA	IOVA	STONY	N.Y.	CORTLAND	PA	FAYETTE
ARIZ	COCHISE	KAN	COBLEY	N.Y.	FULTON	PA	FOULKIN
ARIZ	PINAL	KAN	DOUGLAS	N.Y.	GENESEE	PA	INDIANA
ARK	GARLAND	KAN	LEAVENWORTH	N.Y.	JEFFERSON	PA	LAWRENCE
ARK	MISSISSIPPI	KAN	LYON	N.Y.	ORANGE	PA	LEBANON
ARK	WASHINGTON	KAN	PERO	N.Y.	ST. LAWRENCE	PA	MERCER
CALIF	KINGS	KY	FRANKLIN	N.Y.	STURGEN	PA	NORTHUMBERLAND
CALIF	MADEIRA	KY	HARDIN	N.Y.	TOMPKINS	PA	SCHUYLKILL
CALIF	MERCER	KY	HADISON	N.Y.	ULSTER	R.I.	NEWPORT
CALIF	SAN LUIS ORO	LA	ACADIA	N.Y.	WARREN	S.C.	ANDERSON
CALIF	CUTTER	LA	IMPERIA	N.C.	ALABAMA	S.C.	REAFORTY
CALIF	TULARE	LA	LAFAYETTE	N.C.	CARROLLUS	S.C.	MORRIS
CALIF	YUBA	LA	LINCOLN	N.C.	CLEVELAND	S.C.	YORK
COLO	LARIMER	LA	ST. LANDRY	N.C.	DEDELL	S.C.	SUMTER
COLO	WELD	LA	TANGIPAHOGA	N.C.	DOBSON	TENN	BRADLEY
CONN	LITCHFIELD	LA	VERNON	N.C.	PACHTINGHAM	TENN	COFFEE
CONN	WINDHAM	LA	WASHINGTON	N.C.	POWAY	TENN	MAURY
DEL	KENT	MAINE	KENNEBEC	N.C.	ASHLAND	TENN	MONTEGOMERY
FLA	COLLIER	MAINE	YORK	OHIO	ASHTABULA	TENN	ROANE
FLA	INDIAN RIVER	MD	ALLEGANY	OHIO	ATHENS	TENN	GREGG
FLA	LAKE	MD	FREDERICK	OHIO	COLUMBIANA	TEXAS	HALE
FLA	MANATEE	MD	WASHINGTON	OHIO	COALFORD	TEXAS	HARRISON
FLA	MARIAN	MASS	BARNSTABLE	OHIO	EDIE	TEXAS	HUNT
FLA	MIDDLESEX	MASS	FOULKLIN	OHIO	HANCOCK	TEXAS	JIM WELLS
FLA	OKALOOSA	MICH	BERRIEN	OHIO	HURON	TEXAS	KLEBERG
GA	BALDWIN	MICH	LIVIAWEE	OHIO	LICKING	TEXAS	VICTORIA
GA	SPALDING	MICH	WIDLAND	OHIO	MARION	UTAH	CACHE
GA	WARE	MINN	WOWER	OHIO	ROSS	VA	HEMPY
GA	CANYON	MINN	RICE	OHIO	SANDUSKY	VA	MONTEGOMERY
GA	DE KALB	MINN	STEARNES	OHIO	SCIOTO	VA	PITTSYLVANIA
TLL	WALKER	MINN	WYONNA	OHIO	SENECA	WASH	COLLITZ
TLL	WADSWORTH	MISS	FOREST	OHIO	TUSCARAWAS	WASH	KITSAP
TLL	LA SALLE	MISS	JACKSON	OHIO	WAYNE	WASH	SWAGIT
TLL	WADSWORTH	MISS	WADSWORTH	OKLA	KAY	WASH	THURSTON
TLL	STEPHENSON	MO	CHIEF	OKLA	MUSKOGEE	WASH	WALLA WALLA
TLL	WYATT	NEB	DODGE	OKLA	OSAGE	WASH	UNITMAN
TLL	WYATT	NEB	CHICKASAW	OKLA	OSAGE	WASH	DODGE
TND	WYATT	N.M.	WYATT	OKLA	OSAGE	WISC	FOND DU LAC
TND	WYATT	N.M.	WYATT	OKLA	OSAGE	WISC	JEFFERSON
TND	WYATT	N.M.	WYATT	OKLA	OSAGE	WISC	MAITOVOC
TND	WYATT	N.M.	WYATT	OKLA	OSAGE	WISC	ROCK
TND	WYATT	N.M.	WYATT	OKLA	OSAGE	WISC	SHERIDAN
TND	WYATT	N.M.	WYATT	OKLA	OSAGE	WISC	VALUORTH

HUMAN RESOURCES PROFILE COUNTIES

CAGE 55: NATION. UNRANIZED NONADJACENT

ALA	CAFFE	LA	TEOFRONNE	OKLA	CATFO
ALA	DALE	MAINE	ANDSTOCK	OKLA	GARFIELD
ALA	MUSTON	MAINE	DEWONSON	ORE	JACKSON
ALASKA	FAIRBANKS	MICH	CHADREVA	ORE	COS
ARIZ	COCONINO	MICH	DELYA	ORE	JACKSON
ARIZ	YUMA	MICH	TERRELLA	PA	KLAMATH
ARK	ARLINGTON	MICH	WAGGUYTE	PA	MC KEAN
ARK	COALFIELD	MINN	BLUE EARTH	PA	VFNANCO
ARK	SMITH TPC	MINN	TEBANT	S.C.	FLORENCE
ARK	UNION	MINN	LAKE	S.C.	GREENWOOD
CALIF	BUTTE	MISS	ROLIVAR	S.C.	BROWN
CALIF	MUMBOLDT	MISS	CRANFMA	S.C.	DENNINGTON
CALIF	TYPEPITAL	MISS	JONES	TENN	GIRSON
CALIF	CHASTA	MISS	LAUDEDALE	TENN	MARLEN
COLN	WESA	MISS	LEE	TENN	MADISON
FLA	RAY	MISS	LEFLARE	TENN	SULLIVAN
FLA	ST. LUCIE	MISS	LEFLARE	TENN	WASHINGTON
GA	CLARKE	MISS	LONDES	TEXAS	ANGELIA
GA	FLOYD	MO	WASHINGTON	TEXAS	FRAY
GA	SLYNN	MO	CAPE GIRAROE	TEXAS	MUWAD
GA	LOWYFC	MO	JACPER	TEXAS	LAWAR
GA	TOCUP	MO	MARTON	TEXAS	MACGODACHES
HAWAII	HAWAII	MO	DETTIS	TEXAS	VAL VERNE
HAWAII	BANKYCK	MO	PULASKI	VT	CHITTENDEN
TOHMO	ROYNELVILLE	MONT	LEWIS AND CL	VT	WASHINGTON
TOHMO	NFT SPICE	MONT	MISGULA	VA	ALPHEARLE
TOHMO	TWIN FALLS	NFB	SYLVIA BOW	VA	AUGUSTA
TLL	AGAME	NFB	ADAMS	VA	GPANT
TLL	COLES	NFB	HALL	WASH	GDAYS HARBOR
TLL	JACKSON	NFB	SCOTT BLUFF	WASH	WHATCO
TLL	WC D'NOUSH	N.M.	GOASTON	WASH	MARICON
TLL	WILLIAMSON	N.M.	CHAVES	W.VA	MARTON
TND	MUWAD	N.M.	CURRY	W.VA	MERCER
TOWA	CFRAN GARDN	N.M.	FROY	W.VA	WONGALIA
TOWA	OS WAINES	N.M.	LFA	W.VA	EAU CLAINE
TOWA	LEE	N.M.	OTERN	WISC	MARTON
TOWA	WADSHALL	N.C.	CLINTON	WISC	PORTAGE
TOWA	VADELL	N.C.	CATANTA	WISC	WOOD
TOWA	WBSSTF	N.C.	CHAVEN	WYO	ALBANY
KAN	CAUCARD	N.C.	ENGSCOMRE	WYO	LARAHIF
KAN	GFADY	N.C.	LEVINT	WYO	MATRONA
KAN	MONTGOMERY	N.C.	ONCLON		
KAN	DILFV	N.C.	DITT		
KAN	CALINE	N.C.	WAYNE		
KV	CHISTIAN	N.D.	WILSON		
KV	WC CRACKEN	N.D.	RIPLEIGH		
KV	MARSEN	N.D.	SPAWN CORNS		
LA	ST. MARY	N.D.	WARD		
			WICKINGUM		







[illegible][illegible]

52	JFKING
53	LAVIPO
54	LAUDFNC
55	LINDKIN
56	LINDFTHMD
57	MURRAY
58	PIERCE
59	RAYNOLLO
60	RYTHLE
61	STEPHENS
62	TATNALL
63	TELFALD
64	TIFT
65	TODMRS
66	TOULEN
67	TURNER
68	WASON
69	WASHINGTON
70	WAYNE
71	WILKES
72	WISLAT
73	WU
74	WOLF LAKE
75	WINGHAM
76	WINDHAY
77	WYATT
78	CASSIA
79	CLAFMATED
80	FOAMKLN
81	FOEMNT
82	GODDINS
83	INAHG
84	JEGONY
85	LATAN
86	LEWHI
87	MADISON
88	MINIONNA
89	MYETTF
90	ORVED
91	SHOSMONS
92	WASHINGTON
93	ALEXANDER
94	CARDILL
95	CLAY
96	FFINGHAM
97	FOANKLIN
98	GREENE
99	HAMILTON

ILL	MANCOCK
ILL	MOQUOIS
ILL	JASPER
ILL	JEFFERSON
ILL	LAWRENCE
ILL	LEE
ILL	MASSAC
ILL	PERRY
ILL	PIKE
ILL	RICHLAND
ILL	SALINE
ILL	SCHUYLER
ILL	UNION
ILL	WARREN
ILL	WAYNE
IND	CASS
IND	DAVIES
IND	FAYETTE
IND	JACKSON
IND	JENNINGS
IND	LAWRENCE
IND	MARTIN
IND	MIAMI
IND	ORANGE
IND	PERRY
IND	WARASH
IOVA	ALLAMAKEE
IOVA	APPAYOOSI
IOVA	AUDUBON
IOVA	BUENA VISTA
IOVA	CASSELL
IOVA	CHECKASAM
IOVA	CLAY
IOVA	CALFORD
IOVA	DAVIS
IOVA	DECATUR
IOVA	NICKINSON
IOVA	PAWNE
IOVA	FAYETTE
IOVA	FLOYD
IOVA	FRANKLIN
IOVA	GREEN
IOVA	HAMILTON
IOVA	MARGIN
IOVA	MENDOTA
IOVA	MURADO
IOVA	MURDOCK
IOVA	JEFFERSON



HUMAN RESOURCES PROFILE COUNTIES

CODE 7: HANMPRO- LFSS URRANTIZED NONADJACENT

TOVA	KANSAS	THOMAS	MD	CHESSET	MINN	PENNINGTON
TOVA	KAN	WILSON	MD	TALBOT	MINN	PIESTONE
TOVA	KY	APATO	MD	WICOMICO	MINN	POLK
TOVA	KY	RAPOFF	MD	WRECESTER	MINN	POPE
TOVA	KY	RELL	MICH	ALFRED	MINN	REDWOOD
TOVA	KY	ACOLF	MICH	ALFRED	MINN	SENVILLE
TOVA	KY	CALDWELL	MICH	PARAGA	MINN	DOSEAU
TOVA	KY	CALLOWAY	MICH	CHARLEVOIX	MINN	STEELE
TOVA	KY	CARROLL	MICH	CHENOVAN	MINN	STEVENS
TOVA	KY	COITENDEN	MICH	CLARK	MINN	SWIFT
TOVA	KY	FLOYD	MICH	DICKINSON	MINN	TODD
TOVA	KY	FULTON	MICH	EMFT	MINN	MADENA
TOVA	KY	GRAVES	MICH	ENGELIC	MINN	WASECA
TOVA	KY	GRAYSON	MICH	GRAND TRAVER	MINN	WATSONMAN
TOVA	KY	MARLAN	MICH	HAUGHTON	MISS	YELLOW MEDIC
TOVA	KY	MORRIS	MICH	HUBON	MISS	ADAMS
TOVA	KY	JOHNSON	MICH	ISCO	MISS	ALCORN
TOVA	KY	KNOX	MICH	ISCO	MISS	ATTALA
TOVA	KY	LAQUE	MICH	MACKINAC	MISS	CHICKASAW
TOVA	KY	LAUFEL	MICH	MANISTEE	MISS	CLARKE
TOVA	KY	LEITCHER	MICH	MENOMINEE	MISS	CLAY
TOVA	KY	MARTIN	MICH	OTSEGO	MISS	GRANADA
TOVA	KY	MARSHALL	MICH	PEESQUE ISLE	MISS	HOLMS
TOVA	KY	MASON	MICH	SCHOOLCRAFT	MISS	HUMPHREYS
TOVA	KY	MEAD	MICH	WEXFORD	MISS	ITAMARA
TOVA	KY	MULLENBERG	MICH	WFLTRAMI	MISS	LAFAYETTE
TOVA	KY	PERDY	MINN	WFTON	MISS	LEAKE
TOVA	KY	PIKE	MINN	RIG STONE	MISS	LINCOLN
TOVA	KY	PULASKI	MINN	ROON	MISS	MARION
TOVA	KY	SHAW	MINN	CHIDDEVA	MISS	MORICE
TOVA	KY	TAYLOR	MINN	COTTINGWOOD	MISS	MONTGOMERY
TOVA	KY	WASHINGTON	MINN	COO WING	MISS	NESHOBA
TOVA	KY	WAYNE	MINN	CONGLAS	MISS	WATON
TOVA	KY	WATLEY	MINN	FATRALLY	MISS	NOXURFF
TOVA	LA	CATAMOLLA	MINN	PREETON	MISS	OKTAREMA
TOVA	LA	CONCORDIA	MINN	MURRAON	MISS	PANOLA
TOVA	LA	EAST CARROLL	MINN	JACKSON	MISS	PIKE
TOVA	LA	FRANKLIN	MINN	KANARAC	MISS	PONTOTOC
TOVA	LA	MADISON	MINN	KANDIOWIT	MISS	PONTIAC
TOVA	LA	SCARINE	MINN	KNOCHING	MISS	QUITMAN
TOVA	LA	WIV	MINN	LYON	MISS	SU-FLORNO
TOVA	MAINE	WAVENOX	MINN	MATTIN	MISS	TALLAHATCHIE
TOVA	MAINE	KNOX	MINN	MILLE LACK	MISS	TIPPAN
TOVA	MAINE	WISCATANQUIS	MINN	WOBATON	MISS	UNION
TOVA	MAINE	WAVENOX	MINN	NICOLLET	MISS	WAYNE
TOVA	MAINE	WALON	MINN	NORLES	MISS	WINSTON
TOVA	MAINE	WASHINGTON	MINN	OTTER TAIL	MISS	VALOUSA
TOVA	MAINE	DODCHESER	MD		MO	APAZ









HUMAN RESOURCES PROFILE COUNTIES

CODE 8: Nonmetec. Totally Rural Adjacent

TEXAS	MENDALL	WISC	SAVFIELD
TEXAS	LIVE OAK	WISC	RUPNETT
TEXAS	MARTIN	WISC	BOLK
TEXAS	NEWTON	WISC	TOEMPEALEAU
TEXAS	OLDHAM	WISC	WASHBURN
TEXAS	CAN JACINTA	WISC	WAUCHADA
TEXAS	CHLIFCHP.		
TEXAS	CHACKLEWOOD		
TEXAS	CONFOVELL		
TEXAS	STERLING		
TEXAS	STONFALL		
TEXAS	PADATA		
UTAH	MORGAN		
UTAH	CANDITE		
UTAH	SUMMIT		
VA	AVELTA		
VA	BOUNSWICK		
VA	RICKINGHAM		
VA	CAROLINE		
VA	CHARLOTTE		
VA	CLARK		
VA	CUMBERLAND		
VA	FLOYD		
VA	FLUVANNA		
VA	GILET		
VA	KING AND QUE		
VA	KING GEORGE		
VA	LOUISA		
VA	MATHEWS		
VA	MITCHELL		
VA	MELSON		
VA	NEW KENT		
VA	PATRICK		
VA	STAFFORD		
VA	SURRY		
VA	SUSSEX		
WASH	LINCOLN		
WASH	OFND ORTELL		
WASH	SKAMANIA		
W-VA	RODNEY		
W-VA	CALHOUN		
W-VA	CLAY		
W-VA	LINCOLN		
W-VA	WARRNE		
W-VA	DEACANTS		
W-VA	STITCHIE		
W-VA	ESANF		
W-VA	TYLER		

**CODE 0: counter. Totally Rural**

[illegible]









STANDARD METROPOLITAN STATISTICAL AREAS IN ALPHABETICAL ORDER,  
CONSTITUENT COUNTIES, AND POPULATION RANK AS OF APRIL 1, 1970

SMSA and Constituent Counties	Rank 1970	SMSA and Constituent Counties	Rank 1970
Abilene, Tex. Jones County Taylor County <sup>1</sup>	207	Appleton-Oshkosh, Wis. Outagamie County <sup>1</sup> (Appleton (part)) Calumet County <sup>1</sup> (Appleton (part)) Winnebago County <sup>1</sup> (Oshkosh)	114
Akron, Ohio Portage County Summit County <sup>1</sup>	48	Asheville, N.C. Buncombe County <sup>1</sup>	180
Albany, Ga. Dougherty County <sup>1</sup>	224	Atlanta, Georgia Clayton County Cobb County De Kalb County <sup>1</sup> Fulton County <sup>1</sup> Gwinnett County	20
Albany-Schenectady-Troy, N.Y. Albany County <sup>1</sup> (Albany) Rensselaer County <sup>1</sup> (Troy) Saratoga County Schenectady County <sup>1</sup> (Schenectady)	45	Atlantic City, N.J. Atlantic County <sup>1</sup>	159
Albuquerque, N. Mex. Bernalillo County <sup>1</sup>	96	Augusta, Ga.-S.C. Richmond County, Ga. <sup>1</sup> Aiken County, S.C.	124
Allentown-Bethlehem-Easton, Pa.-N.J. Lehigh County, Pa. <sup>1</sup> (Allentown and Bethlehem (part)) Northampton County, Pa. <sup>1</sup> (Bethlehem (part) and Easton) Warren County, N.J.	58	Austin, Tex. Travis County <sup>1</sup>	103
Altoona, Pa. Blair County <sup>1</sup>	187	Bakersfield, Calif. Kern County <sup>1</sup>	91
Amarillo, Tex. Potter County <sup>1</sup> Randall County <sup>1</sup>	181	Baltimore, Md. Baltimore city Anne Arundel County Baltimore County Carroll County Harford County Howard County	11
Anaheim-Santa Ana-Garden Grove, Calif. Orange County <sup>1</sup>	18	Baton Rouge, La. East Baton Rouge Parish <sup>1</sup>	110
Anderson, Ind. Madison County <sup>1</sup>	185	Bay City, Mich. Bay County <sup>1</sup>	203
Ann Arbor, Mich. Washtenaw County <sup>1</sup>	131		

<sup>1</sup> County in which central city is located.

SMSA and Constituent Counties	Rank 1970	SMSA and Constituent Counties	Rank 1970
Beaumont-Port Arthur-Orange, Tex. Jefferson County <sup>1</sup> (Beaumont and Port Arthur) Orange County <sup>1</sup> (Orange)	95	Buffalo, N.Y. Erie County <sup>1</sup> Niagara County	24
Billings, Mont. Yellowstone County <sup>1</sup>	225	Canton, Ohio Stark County <sup>1</sup>	80
Biloxi-Gulfport, Miss. Harrison County <sup>1</sup>	188	Cedar Rapids, Iowa Linn County <sup>1</sup>	167
Binghamton, N.Y.-Pa. Broome County, N.Y. <sup>1</sup> Tioga County, N.Y. Susquehanna County, Pa.	100	Champaign-Urbana, Ill. Champaign County <sup>1</sup>	166
Birmingham, Ala. Jefferson County <sup>1</sup> Shelby County Walker County	44	Charleston, S.C. Berkeley County Charleston County <sup>1</sup>	99
Bloomington-Normal, Ill. McLean County <sup>1</sup>	215	Charleston, W. Va. Kanawha County <sup>1</sup>	134
Boise City, Idaho Ada County <sup>1</sup>	208	Charlotte, N.C. Mecklenburg County <sup>1</sup> Union County	73
Boston, Mass. Essex County (part) Middlesex County (part) Norfolk County (part) Plymouth County (part) Suffolk County (part) <sup>1</sup>	8	Chattanooga, Tenn.-Ga. Hamilton County, Tenn. <sup>1</sup> Walker County, Ga.	97
Bridgeport, Conn. Fairfield County (part) <sup>1</sup> New Haven County (part)	76	Chicago, Ill. Cook County <sup>1</sup> Du Page County <sup>1</sup> Kane County Lake County Mc Henry County Will County	3
Bristol, Conn. Hartford County (part) <sup>1</sup> Litchfield County (part)	240	Cincinnati, Ohio-Ky.-Ind. Clermont County, Ohio Hamilton County, Ohio <sup>1</sup> Warren County, Ohio Boone County, Ky. Campbell County, Ky. Kenton County, Ky. Dearborn County, Ind.	21
Brockton, Mass. Bristol County (part) Norfolk County (part) Plymouth County (part) <sup>1</sup>	151	Cleveland, Ohio Cuyahoga County <sup>1</sup> Geauga County Lake County Medina County	12
Brownsville-Harlingen-San Benito, Tex. Cameron County <sup>1</sup>	184		
Bryan-College Station, Tex. Brazos County <sup>1</sup>	242		

<sup>1</sup> County in which central city is located.

SMSA and Constituent Counties	Rank 1970	SMSA and Constituent Counties	Rank 1970
Colorado Springs, Colo. El Paso County <sup>1</sup>	129	Decatur, Ill. Macon County <sup>1</sup>	19
Gadsden, Ala. Etowah County <sup>1</sup>	221	Denver, Colo. Adams County Arapahoe County Boulder County Denver County <sup>1</sup> Jefferson County	27
Columbia, Mo. Boone County <sup>1</sup>	231	Des Moines, Iowa Polk County <sup>1</sup>	109
Columbia, S.C. Lexington County Richland County <sup>1</sup>	93	Detroit, Mich. Macomb County Oakland County Wayne County <sup>1</sup>	5
Columbus, Ga.-Ala. Chattahoochee County, Ga. Muscogee County, Ga. <sup>1</sup> Russell County, Ala.	35	Dubuque, Iowa Dubuque County <sup>1</sup>	223
Columbus, Ohio Delaware County Franklin County <sup>1</sup> Pickaway County	129	Duluth-Superior, Minn.-Wis. St. Louis County, Minn. <sup>1</sup> (Duluth) Douglas County, Wis. <sup>1</sup> (Superior)	117
Corpus Christi, Tex. Nueces County <sup>1</sup> San Patricio County	111	Durham, N.C. Durham County <sup>1</sup> Orange County	150
Dallas, Tex. Collin County Dallas County <sup>1</sup> Denton County Ellis County Kaufman County Rockwall County	16	El Paso, Tex. El Paso County <sup>1</sup>	82
Danbury, Conn. Fairfield County (part) <sup>1</sup>	235	Erie, Pa. Erie County <sup>1</sup>	119
Davenport-Rock Island-Moline, Iowa-Ill. Scott County, Iowa <sup>1</sup> (Davenport) Rock Island County, Ill. <sup>1</sup> (Moline and Rock Island) Henry County, Ill.	81	Eugene, Oreg. Lane County <sup>1</sup>	139
Dayton, Ohio Greene County Miami County Montgomery County <sup>1</sup> Preble County	39	Evansville, Ind.-Ky. Vanderburgh County, Ind. <sup>1</sup> Warrick County, Ind. Henderson County, Ky.	132
		Fall River, Mass.-R.I. Bristol County, Mass. (part) <sup>1</sup> Newport County, R.I. (part)	176
		Fargo-Moorhead, N. Dak.-Minn. Cass County, N. Dak. <sup>1</sup> (Fargo) Clay County, Minn. <sup>1</sup> (Moorhead)	199

<sup>1</sup> County in which central city is located.

SMSA and Constituent Counties	Rank 1970	SMSA and Constituent Counties	Rank 1970
Fayetteville, N.C. Cumberland County <sup>1</sup>	141	Greensboro-Winston-Salem-High Point, N.C. Forsyth County <sup>1</sup> (Winston-Salem) Guilford County <sup>1</sup> (Greensboro and High Point) Randolph County Yadkin County	56
Fitchburg-Leominster, Mass. Middlesex County (part) Worcester County (part) <sup>1</sup>	218		
Flint, Mich. Genesee County <sup>1</sup> Lapeer County	67	Greenville, S.C. Greenville County <sup>1</sup> Pickens County	101
Fort Lauderdale-Hollywood, Fla. Broward County <sup>1</sup>	54	Hamilton-Middletown, Ohio Butler County <sup>1</sup>	137
Fort Smith, Ark.-Okla. Sebastian County, Ark. <sup>1</sup> Crawford County, Ark. Le Flore County, Okla. Sequoyah County, Okla.	169	Harrisburg, Pa. Cumberland County Dauphin County <sup>1</sup> Perry County	72
Fort Wayne, Ind. Allen County <sup>1</sup>	112	Hartford, Conn. Hartford County (part) <sup>1</sup> Middlesex County Tolland County	49
Fort Worth, Tex. Johnson County Tarrant County <sup>1</sup>	43	Honolulu, Hawaii Honolulu County <sup>1</sup>	53
Fresno, Calif. Fresno County <sup>1</sup>	70	Houston, Texas Brazoria County Fort Bend County Harris County <sup>1</sup> Liberty County Montgomery County	13
Gainesville, Fla. Alachua County <sup>1</sup>	214		
Galveston-Texas City, Tex. Galveston County <sup>1</sup>	163	Huntington-Ashland, W. Va.-Ky.-Ohio Cabell County, W. Va. <sup>1</sup> (Huntington (part)) Wayne County (Huntington (part)) Boyd County, Ky. <sup>1</sup> (Ashland) Lawrence County, Ohio	123
Gary-Hammond-East Chicago, Ind. Lake County <sup>1</sup> Porter County	52	Huntsville, Ala. Limestone County Madison County <sup>1</sup>	136
Grand Rapids, Mich. Kent County <sup>1</sup> Ottawa County	61	Indianapolis, Ind. Boone County Hamilton County Hancock County Hendricks County Johnson County Marion County <sup>1</sup>	29
Great Falls, Mont. Cascade County <sup>1</sup>	230		
Green Bay, Wis. Brown County <sup>1</sup>	170	Morgan County Shelby County	

<sup>1</sup> County in which central city is located.

SMSA and Constituent Counties	Rank 1970	SMSA and Constituent Counties	Rank 1970
Jackson, Mich. Jackson County <sup>1</sup>	182	Lansing, Mich. Clinton County Eaton County Ingham County <sup>1</sup>	77
Jackson, Miss. Hinds County <sup>1</sup> Rankin County	121	Laredo, Tex. Webb County <sup>1</sup>	236
Jacksonville, Fla. Duval County <sup>1</sup>	64	Las Vegas, Nev. Clark County <sup>1</sup>	115
Jersey City, N.J. Hudson County <sup>1</sup>	55	Lawrence-Haverhill, Mass.-N.H. Essex County, Mass. (part) <sup>1</sup> Rockingham County, N.H. (part)	133
Johnstown, Pa. Cambria County <sup>1</sup> Somerset County	120	Lawton, Okla. Comanche County <sup>1</sup>	212
Kalamazoo, Mich. Kalamazoo County <sup>1</sup>	147	Lewiston-Auburn, Maine Androscoggin County (part) <sup>1</sup>	237
Kansas City, Mo.-Kans. Cass County, Mo. Clay County, Mo. <sup>1</sup> (Kansas City (part)) Jackson County, Mo. <sup>1</sup> (Kansas City (part)) Platte County, Mo. Johnson County, Kans. Wyandotte County, Kans.	26	Lexington, Ky. Fayette County <sup>1</sup>	160
Kenosha, Wis. Kenosha County <sup>1</sup>	202	Lima, Ohio Allen County <sup>1</sup> Putnam County Van Wert County	181
Knoxville, Tenn. Anderson County Blount County Knox County <sup>1</sup>	74	Lincoln, Nebr. Lancaster County <sup>1</sup>	164
La Crosse, Wis. La Crosse County <sup>1</sup>	232	Little Rock-North Little Rock, Ark. Pulaski County <sup>1</sup> Saline County	92
Lafayette, La. Lafayette Parish <sup>1</sup>	209	Lorain-Elyria, Ohio Lorain County	122
Lafayette-West Lafayette, Ind. Tippacanoe County <sup>1</sup>	210	Los Angeles-Long Beach, Calif. Los Angeles County <sup>1</sup>	2
Lake Charles, La. Calcasieu Parish <sup>1</sup>	178	Louisville, Ky.-Ind. Jefferson County, Ky. <sup>1</sup> Clark County, Ind. Floyd County, Ind.	40
Lancaster, Pa. Lancaster County <sup>1</sup>	94	Lowell, Mass. Middlesex County (part) <sup>1</sup>	140

<sup>1</sup> County in which central city is located.

SMSA and Constituent Counties	Rank 1970	SMSA and Constituent Counties	Rank 1970
Lubbock, Tex. Lubbock County <sup>1</sup>	157	Mobile, Ala. Baldwin County Mobile County <sup>1</sup>	78
Lynchburg, Va. Lynchburg city Amherst County Campbell County	196	Modesto, Calif. Stanislaus County <sup>1</sup>	149
Macon, Ga. Bibb County <sup>1</sup> Houston County	145	Monroe, La. Ouachita Parish <sup>1</sup>	206
Madison, Wis. Dane County <sup>1</sup>	106	Montgomery, Ala. Elmore County Montgomery County <sup>1</sup>	148
Manchester, N.H. Hillsborough County (part) <sup>1</sup> Merrimack County (part)	211	Muncie, Ind. Delaware County <sup>1</sup>	191
Mansfield, Ohio Richland County <sup>1</sup>	190	Muskegon-Muskegon Heights, Mich. Muskegon County <sup>1</sup>	171
Mc Allen-Pharr-Edinburg, Tex. Hidalgo County <sup>1</sup>	155	Nashua, N.H. Hillsborough County (part) <sup>1</sup>	239
Memphis, Tenn.-Ark. Shelby County, Tenn. <sup>1</sup> Crittenden County, Ark.	42	Nashville, Tenn. Davidson County <sup>1</sup> Sumner County Wilson County	59
Meriden, Conn. New Haven County (part) <sup>1</sup>	243	New Bedford, Mass. Bristol County (part) <sup>1</sup> Plymouth County (part)	175
Miami, Fla. Dade County <sup>1</sup>	25	New Britain, Conn. Hartford County (part) <sup>1</sup>	179
Midland, Tex. Midland County <sup>1</sup>	241	New Haven, Conn. New Haven County (part) <sup>1</sup>	83
Milwaukee, Wis. Milwaukee County <sup>1</sup> Ozaukee County Washington County Waukesha County	19	New London-Groton-Norwich, Conn. New London County (part) <sup>1</sup>	143
Minneapolis-St. Paul, Minn. Anoka County Dakota County Hennepin County <sup>1</sup> Ramsey County <sup>1</sup> Washington County	15	New Orleans, La. Jefferson Parish Orleans Parish <sup>1</sup> St. Bernard Parish St. Tammany Parish	31

<sup>1</sup> County in which central city is located.



SMSA and Constituent Counties	Rank 1970	SMSA and Constituent Counties	Rank 1970
New York, N.Y.	1	Oxnard-Ventura, Calif.	79
New York City		Ventura County <sup>1</sup>	
Bronx County <sup>1</sup>			
Kings County <sup>1</sup>		Owensboro, Ky.	234
New York County <sup>1</sup>		Davies County <sup>1</sup>	
Queens County <sup>1</sup>			
Richmond County <sup>1</sup>		Paterson-Clifton-Passaic, N.J.	22
Nassau County		Bergen County	
Rockland County		Passaic County <sup>1</sup>	
Suffolk County			
Westchester County		Pensacola, Fla.	127
Newark, N.J.	14	Escambia County <sup>1</sup>	
Essex County <sup>1</sup>		Santa Rosa County	
Morris County			
Union County		Peoria, Ill. <sup>1</sup>	88
Newport News-Hampton, Va.	105	Peoria County <sup>1</sup>	
Hampton city <sup>1</sup>		Tazewell County	
Newport News city <sup>1</sup>		Woodford County	
York County			
Norfolk-Portsmouth, Va.	47	Petersburg-Colonial Heights, Va.	192
Chesapeake city		Colonial Heights city <sup>1</sup>	
Norfolk city <sup>1</sup>		Hopewell city	
Portsmouth city <sup>1</sup>		Petersburg city <sup>1</sup>	
Virginia Beach city		Dinwiddie County	
Norwalk, Conn.	200	Prince George County	
Fairfield County <sup>1</sup> (part)			
Odessa, Tex.	222	Philadelphia, Pa.-N.J.	4
Ector County <sup>1</sup>		Bucks County, Pa.	
Ogden, Utah	104	Chester County, Pa.	
Weber County <sup>1</sup>		Delaware County, Pa.	
Oklahoma City, Okla.	50	Montgomery County, Pa.	
Canadian County <sup>1</sup> (part)		Philadelphia County, Pa. <sup>1</sup>	
Cleveland County <sup>1</sup> (part)		Burlington County, N.J.	
Oklahoma County <sup>1</sup> (part)		Camden County, N.J.	
Omaha, Nebr.-Iowa	60	Gloucester County, N.J.	
Douglas County, Nebr. <sup>1</sup>		Phoenix, Ariz.	34
Sarpy County, Nebr.		Maricopa County <sup>1</sup>	
Pottawattamie County, Iowa			
Orlando, Fla.	69	Pine Bluff, Ark.	227
Orange County <sup>1</sup>		Jefferson County <sup>1</sup>	
Seminole County			
		Pittsburgh, Pa.	9
		Allegheny County <sup>1</sup>	
		Beaver County	
		Washington County	
		Westmoreland County	
		Pittsfield, Mass.	233
		Berkshire County (part) <sup>1</sup>	

<sup>1</sup> County in which central city is located.

SMSA and Constituent Counties	Rank 1970	SMSA and Constituent Counties	Rank 1970
Portland, Maine	183	St. Louis, Mo.-Ill.	10
Cumberland County (part) <sup>1</sup>		St. Louis city, Mo. <sup>1</sup>	
Portland, Oreg.-Wash.	33	Franklin County, Mo.	
Clackamas County, Oreg. <sup>1</sup>		Jefferson County, Mo.	
Multnomah County, Oreg. <sup>1</sup>		St. Charles County, Mo.	
Washington County, Oreg.		St. Louis County, Mo.	
Clark County, Wash.		Madison County, Ill.	
		St. Clair County, Ill.	
Providence-Pawtucket-Warwick, R.I.-Mass.	36	Salt Lake City, Utah	57
Bristol County, R.I.		Davis County	
Kent County, R.I. (part) <sup>1</sup>		Salt Lake City	
Newport County, R.I. (part)		San Angelo, Tex.	238
Providence County, R.I. (part) <sup>1</sup>		Tom Green County <sup>1</sup>	
Washington County, R.I. (part)			
Bristol County, Mass. (part)		San Antonio, Tex.	38
Norfolk County, Mass. (part)		Bexar County <sup>1</sup>	
Worcester County, Mass. (part)		Guadalupe County	
Provo-Orem, Utah	186	San Bernardino-Riverside-Ontario, Calif.	28
Utah County <sup>1</sup>		Riverside County <sup>1</sup>	
Pueblo, Colo.	201	San Bernardino County	
Pueblo County <sup>1</sup>		San Diego, Calif.	23
Racine, Wis.	162	San Diego County <sup>1</sup>	
Racine County <sup>1</sup>		Rochester, Minn.	228
Raleigh, N.C.	135	Olmsted County <sup>1</sup>	
Wake County <sup>1</sup>		Rochester, N.Y.	37
Reading, Pa.	102	Livingston County	
Berks County <sup>1</sup>		Monroe County <sup>1</sup>	
Reno, Nev.	198	Orleans County	
Washoe County <sup>1</sup>		Wayne County	
Richmond, Va.	65	Rockford, Ill.	116
Richmond city <sup>1</sup>		Boone County	
Chesterfield County		Winnebago County <sup>1</sup>	
Hanover County		Sacramento, Calif.	41
Henrico County		Placer County	
Roanoke, Va.	156	Sacramento County <sup>1</sup>	
Roanoke city <sup>1</sup>		Yolo County	
Roanoke County		Saginaw, Mich.	138
		Saginaw County <sup>1</sup>	
		Salem, Oregon	153
		Marion County <sup>1</sup>	
		Polk County	

<sup>1</sup> County in which central city is located.

SMSA and Constituent Counties	Rank 1970	SMSA and Constituent Counties	Rank 1970
Salinas-Monterey, Calif. Monterey County <sup>1</sup>	125	Spokane, Wash. Spokane County <sup>1</sup>	108
St. Joseph, Mo. Buchanan County <sup>1</sup>	226	Springfield, Ill. Sangamon County <sup>1</sup>	168
San Francisco, Calif. Alameda County <sup>1</sup> Contra Costa County Marin County San Francisco County <sup>1</sup> San Mateo County	6	Springfield, Mo. Greene County <sup>1</sup>	174
San Jose, Calif. Santa Clara County <sup>1</sup>	30	Springfield, Ohio Clark County <sup>1</sup>	172
Santa Barbara, Calif. Santa Barbara County <sup>1</sup>	118	Springfield-Chicopee-Holyoke, Mass. Conn. Hampden County, Mass. (part) <sup>1</sup> Hampshire County, Mass. (part) Worcester County, Mass. (part) Tolland County, Conn. (part)	63
Santa Rosa, Calif. Sonoma County <sup>1</sup>	146	Stamford, Conn. Fairfield County (part) <sup>1</sup>	144
Savannah, Ga. Chatham County <sup>1</sup>	152	Steubenville-Weirton, Ohio-W. Va. Jefferson County, Ohio <sup>1</sup> Brooke County, W. Va. <sup>1</sup> Hancock County, W. Va. <sup>1</sup>	165
Scranton, Pa. Lackawanna County <sup>1</sup>	130	Stockton, Calif. San Joaquin County <sup>1</sup>	107
Seattle-Everett, Wash. King County <sup>1</sup> Snohomish County <sup>1</sup>	17	Syracuse, N.Y. Madison County Onondaga County <sup>1</sup> Oswego County	51
Sherman-Denison, Texas Grayson County <sup>1</sup>	229	Tacoma, Wash. Pierce County <sup>1</sup>	71
Shreveport, La. Bossier Parish <sup>1</sup> Caddo Parish <sup>1</sup>	104	Tallahassee, Fla. Leon County <sup>1</sup>	216
Sioux City, Iowa-Nebr. Woodbury County, Iowa <sup>1</sup> Dakota County, Nebr.	204	Tampa-St. Petersburg, Fla. Hillsborough County <sup>1</sup> Pinellas County <sup>1</sup>	32
Sioux Falls, S. Dak. Minnehaha County <sup>1</sup>	220	Terre Haute, Ind. Clay County Sullivan County Vermillion County Vigo County <sup>1</sup>	158
South Bend, Ind. St. Joseph County <sup>1</sup> Marshall County	113		

<sup>1</sup> County in which central city is located.

SMSA and Constituent Counties	Rank 1970	SMSA and Constituent Counties	Rank 1970
Texarkana, Tex.-Ark.	217	Washington, D.C.-Md.-Va.-Con.	
Bowie County, Tex. <sup>1</sup>		Arlington County, Va.	
Miller County, Ark. <sup>1</sup>		Fairfax County, Va.	
Toledo, Ohio-Mich.	46	Loudoun County, Va.	
Lucas County, Ohio <sup>1</sup>		Prince William County, Va.	
Wood County, Ohio		Waterbury, Conn.	142
Monroe County, Mich.		Litchfield County (part)	
Topeka, Kans.	173	New Haven County (part) <sup>1</sup>	
Shawnee County <sup>1</sup>		Waterloo, Iowa	189
Trenton, N.J.	98	Black Hawk County <sup>1</sup>	
Mercer County <sup>1</sup>		West Palm Beach, Fla.	85
Tucson, Ariz.	84	Palm Beach County <sup>1</sup>	
Pima County <sup>1</sup>	....	Wheeling, W. Va.-Ohio	154
Tulsa, Okla.	68	Marshall County, W. Va.	
Creek County		Ohio County, W. Va. <sup>1</sup>	
Osage County		Belmont County, Ohio	
Tulsa County <sup>1</sup>		Wichita, Kans.	75
Tuscaloosa, Ala.	205	Butler County	
Tuscaloosa County <sup>1</sup>		Sedgwick County <sup>1</sup>	
Tyler, Texas	219	Wichita Falls, Tex.	193
Smith County <sup>1</sup>		Archer County	
Utica-Rome, N.Y.	89	Wichita County <sup>1</sup>	
Herkimer County		Wilkes-Barre-Hazleton, Pa.	87
Oneida County <sup>1</sup>		Luzerne County <sup>1</sup>	
Vallejo-Napa, Calif.	126	Wilmington, Del.-N.J.-Md.	66
Napa County <sup>1</sup>		New Castle County, Del. <sup>1</sup>	
Solano County <sup>1</sup>		Salem County, N.J.	
Vineland-Millville-Bridgeton, N.J.	197	Cecil County, Md.	
Cumberland County <sup>1</sup>		Wilmington, N.C.	213
Waco, Texas	177	Brunswick County	
Mc Lennan County <sup>1</sup>		New Hanover County <sup>1</sup>	
Washington, D.C.-Md.-Va.	7	Worcester, Mass.	86
District of Columbia <sup>1</sup>		Worcester County (part) <sup>1</sup>	
Montgomery County, Md.		York, Pa.	90
Prince Georges County, Md.		Adams County	
Alexandria city, Va.		York County <sup>1</sup>	
Fairfax city, Va.		Youngstown-Warren, Ohio	62
Falls Church city, Va.		Mahoning County <sup>1</sup>	
		Trumbull County <sup>1</sup>	

<sup>1</sup> County in which central city is located.

Appendix I

INITIAL APPROXIMATION OF COSTS FOR BASIC  
BATTERY TESTING PER SITE<sup>1</sup>

Below are listed estimates of costs involved in setting up an evaluation site and administering the basic battery of dependent variable measures. Omitted are costs involved in the collection of independent variable data, in the administration of measures recommended for subsamples of children (e.g., physical examination, serum albumin for Native Americans) and focused studies, and in contractor overhead. The estimates provided do, however, include tests administered in different versions to different subpopulations of the total sample (e.g., a Spanish-speaking version of CIRCUS measures).

PRETEST -- HEAD START YEAR

(Assumes inclusion of optional pretesting of health battery.)

<u>Item</u>	<u>Estimate (\$)</u>
I. Testing space (rented trailer, independent of classroom) \$300/week x 3 weeks testing <sup>2</sup> .....	900
II. Site coordinator (to manage training of testers, start-up of sites, and actual testing) A. 5 days training x \$30/day .....	150
B. 40 days work x \$30/day .....	1200
III. Clerical assistant to site coordinator (to schedule subjects, etc.) \$2.20/hour x 40 hours/week x 8 weeks .....	704
IV. Basic battery testers A. CIRCUS - 2 testers, each specializing in one-half of the six group measures 1. 2 days training x \$20/day x 2 testers .....	80
2. 9 days testing (six days at 5 subjects a day and three make-up days) x \$3.50/hour x 5 hours/day x 2 testers .....	315

<sup>1</sup>Many of the numbers from which estimates are derived are based on guidelines resulting from the field experience of the Stanford Research Institute in its study of Head Start Planned Variation.

<sup>2</sup>Two weeks scheduled testing and one week make-up testing.

B. Ravens Matrices - 1 tester	
1. 2 days training x \$20/day .....	40
2. 9 days testing x \$3.50/hour x 5 hours/day .....	158
C. Health - 1 dentist, 2 testers	
1. Dentist: \$11.50/child x 60 children per site* .....	690
2. Testers:	
a. 5 days training x \$20/day x 2 testers .....	200
b. 9 days testing x \$3.50/hour x 5 hours/day x 2 testers .....	315
3. Hematocrit processing: \$2.60/child x 60 children per site <sup>1</sup> .....	156
V. Parent honorarium (to cover babysitting, trans- portation costs; for control parents only) \$5/session x 30 parent x 4 sessions .....	600
VI. Test Materials	
A. CIRCUS: (6 tests x \$3.75/10 tests + 1 test x \$1.00/10 tests) x 60 children per site .....	141
B. Ravens: \$14.50/25 tests x 60 children per site .....	35
C. Audiometer: too expensive to rent or buy; probably can borrow from community center or school for token fee .....	25
<b>TOTAL</b>	<b>\$5589</b>

#### POST TEST -- HEAD START YEAR

Same items as listed in I through VI above; however, training costs would be reduced by approximately one-half because of the likely retention of some testers and site coordinator.

**TOTAL** **\$5354**

#### SOCIAL-EMOTIONAL BATTERY -- FIRST SCHOOL YEAR

I. Testing Space (rented trailer, independent of classroom)	
\$300/week x 3 weeks testing .....	900
II. Site coordinator (to manage training of testers, start-up of sites, and actual testing)	
A. 5 days training x \$30/day .....	150
B. 40 days work x \$30/day .....	1200

<sup>1</sup>Based on estimates from Health Start data and the experiences of the Early and Periodic Screening, Diagnosis, and Treatment programs.

III. Clerical assistant to site coordinator (to schedule subjects, etc.)	
\$2.20/hour x 40 hours/week x 8 weeks .....	704
IV. Basic Battery testers	
A. Observational measures - w observers	
1. 7 days training x \$18/day x 2 observers .....	252
2. 7 days observing (1/2 hour/child; 2 children/class/day; 4 classes/day; 8 children/day/observer; 4 days initial, 3 days make-up x 2 observers .....	252
B. Teacher responses - 1 tester to instruct teachers in Q-sort	
1. Teacher honorarium (for 5 measures and approximately 4 hours of time): \$20/teacher x 4 classes .....	80
2. Tester training: 2 days x \$20/day .....	40
3. Tester time: \$3.50/hour x 4 hours .....	14
C. Parent (and others) measures - 2 interviewers	
1. 5 days training x \$20/day x 2 interviewers .....	200
2. 15 days interviewing x \$20/day x 2 interviewers .....	600
D. Subject measures - 2 interviewers	
1. 5 days training x \$20/day x 2 interviewers .....	200
2. 12 days interviewing x \$20/day x 2 interviewers .....	<u>600</u>
TOTAL	\$5192